Graph-Based Methods for Forecasting Realized Covariances

Chao Zhang, Xingyue Pu, Mihai Cucuringu, Xiaowen Dong

Journal of Financial Econometrics, 2025, **23(2)**, nbae026
https://doi.org/10.1093/jjfinec/nbae026
Article

OXFORD

Graph-Based Methods for Forecasting Realized Covariances

Chao Zhang [1,2], Xingyue Pu [2,3], Mihai Cucuringu [2,4,5] and Xiaowen Dong [2,3]

[1]FinTech Thrust, HKUST(GZ), Guangzhou, China
[2]Oxford-Man Institute of Quantitative Finance, University of Oxford, Oxford, UK
[3]Department of Engineering Science, University of Oxford, Oxford, UK
[4]Department of Statistics, University of Oxford, Oxford, UK
[5]The Alan Turing Institute, London, UK

Address correspondence to Chao Zhang, FinTech Thrust, HKUST(GZ), Guangzhou, China or e-mail: chaoz@hkust-gz.edu.cn.

Abstract

We forecast the realized covariance matrix of asset returns in the U.S. equity market by exploiting the predictive information of graphs in volatility and correlation. Specifically, we augment the Heterogeneous Autoregressive model via neighborhood aggregation on these graphs. Our proposed method allows for the modeling of interdependence in volatility (also known as spillover effect) and correlation, while maintaining parsimony and interpretability. We explore various graph construction methods, including sector membership and graphical LASSO (for modeling volatility), and line graph (for modeling correlation). The results generally suggest that the augmented model incorporating graph information yields both statistically and economically significant improvements for out-of-sample performance over the traditional models. Such improvements remain significant over horizons up to 1 month ahead, but decay in time. The robustness tests demonstrate that the forecast improvements are obtained consistently over the different out-of-sample sub-periods and are insensitive to measurement errors of volatilities.

Keywords: realized covariance, HAR, graphical LASSO, line graph, graph learning

JEL classifications: C31, C53, C58, G17

Covariance matrix estimation is of particular importance in areas such as risk management, asset pricing, and portfolio optimization. More recently, the increased availability of high-frequency intraday asset prices has motivated researchers to model and forecast the nonparametric and ex-post Realized Covariance (RC) measures constructed from intraday data (see Andersen et al. 2003; Andersen, Bollerslev, and Meddahi 2011; Barndorff-Nielsen et al. 2011). It has been extensively documented in the literature that the high-frequency-based RCs are superior to those constructed from low-frequency return data, for practical investment and portfolio allocation decisions (e.g., Hautsch, Kyj, and Malec 2015; Bollerslev, Patton, and Quaedvlieg 2018).

This work is supported by the Guangzhou-HKUST(GZ) Joint Funding Program (No. 2024A03J0630). The authors also thank the Oxford Suzhou Centre for Advanced Research for providing computing facilities and the Oxford-Man Institute of Quantitative Finance for support

Received: November 11, 2022. **Revised:** May 28, 2024. **Editorial decision:** October 01, 2024. **Accepted:** October 17, 2024

There is a large volume of studies dedicated to forecasting RC matrices. Chiriac and Voev (2011) extended the univariate Heterogeneous Autoregressive (HAR) model of Corsi (2009), to a multivariate setting via the *Cholesky decomposition* on the covariance matrix.[1] Inspired by the dynamic conditional correlation (DCC) model of Engle (2002), an alternative approach to forecasting RCs is based on the decomposition of the covariance matrix into the volatilities and correlations components, also known as the *DRD decomposition*.[2] Bollerslev, Patton, and Quaedvlieg (2018) reported that covariance forecasts derived from the DRD decomposition, known as the HAR-DRD model, generally exhibit superior predictive performance compared to those derived from the Cholesky decomposition.

This research presents novel empirical assessments of the advantages of modeling RC matrices through variances and correlations. The existing HAR-DRD model is augmented by the inclusion of **graph**-based information. In financial markets, the fluctuations in an individual asset's prices are influenced not only by its own characteristics but also by its interconnectedness with other assets. Specifically, we utilize graphs as the mathematical model to represent different types of interdependence in variances and correlations, separately. In the context of volatility modeling, individual assets are represented as nodes, while edges connecting nodes denote the existence of dependency between their corresponding volatilities. Within the graph pertaining to correlation modeling, each node is a correlation pair of two assets and an edge connecting two nodes symbolizes the relationship between distinct pairs. Further details regarding relevant graph construction and estimation methodologies are briefly reviewed in Section 2.4.1.

The perspective of representing interdependence as a graph allows us to adopt the modeling tools tailored for the analysis of graph-structured data. In particular, we employ the operation of *neighborhood aggregation* to generate new predictors for every underlying asset, which are subsequently integrated into the standard HAR model to enhance the accuracy of volatility forecasting. Analogously, comparable predictors can be generated for each pair of assets to facilitate correlation forecasting. The aggregated predictor encapsulates the influence of the connected assets (resp. pairs of assets) on a specific target asset (resp. a specific pair correlation).[3]

In parallel, previous studies have undertaken a comprehensive examination of volatility spillover effects,[4] revealing potential opportunities for improved volatility forecasting. Conventional methods for capturing spillover effects, including multivariate generalized autoregressive conditional heteroskedasticity and vector autoregression (VAR), may reveal important information about the interactions within cross-sectional volatilities, but encounter challenges in producing reliable out-of-sample forecasts due to the curse of dimensionality (Callot, Kock, and Medeiros 2017). On the other hand, the realm of correlation forecasting, especially considering the interdependence among pairs of correlations, has received relatively scant attention within the existing body of literature. For instance, there remains a lack of clarity regarding the extent and nature of the reciprocal influences between correlation pairs.

To the best of our knowledge, the conceptualization of such forms of interdependence using graph-based models has not been extensively explored in the research literature. The

1 This model was also used, with some variations, in other papers, for example, Symitsi et al. (2018); Bucci (2020b); Fiszeder and Orzeszko (2021); Bollerslev et al. (2022).

2 For additional instances, refer to Lee and Long (2009); Oh and Patton (2016); Vassallo, Buccheri, and Corsi (2021).

3 We adopt a straightforward method for neighborhood aggregation, that is computing the average value of connected neighbors. It is worth noting that future investigations could explore alternative permutation invariant operations, such as maximum and median.

4 In our article, volatility spillover effects refer to the transmission of volatility from one market, asset, or sector to another (Baele 2005; Diebold and Yilmaz 2012; Bonato, Caporin, and Ranaldo 2013), which may result from market interdependencies, shared information flows, etc.

approach of neighborhood aggregation in the context of graphs implies a specific impact parameter of the aggregated lagged variances (or correlations) of connected neighbors on the subsequent variance (or correlation) of interest. Remarkably, the application of neighborhood aggregation maintains predictor dimensionality at parity with the original attributes, thereby endowing the proposed models with heightened informativeness compared to the extant HAR-DRD framework, while retaining a parsimonious structure.

The main contributions of this article are summarized as follows. First, we put forward a methodology to model and forecast RC matrices by leveraging the graph information. Our method may shed new light on the dynamic interdependence of realized volatility and correlation, through the lens of neighborhood aggregation. Second, we apply our method to forecasting RC matrices over the daily, weekly, and monthly future horizons, in order to study the variation in the predictive power of graphs across time. In the end, we examine the stability and robustness of our model through comprehensive experiments.

More specifically, using data for the components of the Dow Jones Industrial Average, spanning the period from July 2007 to June 2021, we find that graph information of volatility and correlation can be used to improve the forecasts of RC matrices. Our in-sample results first show that lagged daily volatility is the most important source for future volatility forecasting, while the correlation model attributes the most importance to the weekly component. This highlights the different dependencies in volatilities and correlations. Moreover, our in-sample analysis shows that graph-based predictors for volatility (resp. correlation) modeling are highly significant, and including these graph components in the traditional models can substantially improve the in-sample fit accuracy of RC matrices.

By performing the out-of-sample tests in a rolling window setup, our analysis provides strong evidence that the graph-based model yields superior 1-day-ahead forecasts over a strong set of traditional baselines. The Model Confidence Set (MCS) test (Hansen, Lunde, and Nason 2011) indicates that the augmented model with a data-driven graph determined by Graphical LASSO (GL; Friedman, Hastie, and Tibshirani 2008) is consistently included in the subset of best forecasting models irrespective of the loss functions. The graph analysis highlights the time-varying nature of graph effects both in volatilities and correlations. We also investigate the economic benefits of covariance forecasts from different models via the global minimum variance portfolio (GMVP). The results show that portfolios employing graph information are able to generate significantly lower out-of-sample variance compared to the traditional models without graph information or the naive equally weighted portfolio. The forecast improvements over longer horizons (1-week, 1-month) remain significant, but decay as the prediction horizon gets longer. We conclude the article by several robustness tests, which demonstrate the forecast improvements are experienced consistently over the different out-of-sample sub-periods and are insensitive to measurement errors of volatilities (Bollerslev, Patton, and Quaedvlieg 2016, 2018).

The rest of the article is structured as follows. We begin with Section 1 by reviewing some closely related literature. In Section 2, we first introduce the estimation of RC matrices from high-frequency data and the baseline models, then present the background of graph and propose a new methodology based on graphs for modeling and forecasting RC matrices. Section 3 provides a description of the dataset and evaluation criteria, followed by the in-sample analysis and out-of-sample results. In Section 4, we perform several robustness tests. Finally, we conclude our analysis in Section 5 and highlight potential future research directions.

1 Related Literature

Numerous contributions have been made to the literature on forecasting daily covariance matrices. Notably, the BEKK-GARCH model introduced by Engle and Kroner (1995) and the DCC-type models by Engle (2002) are frequently used for this purpose. The former dominantly suffers from the curse of dimensionality, while the latter allows modeling and estimating the parameters of the variances and correlations separately. The scalar DCC-type model is relatively more flexible and usually provides a better fit and forecasting quality than the scalar BEKK-type model, underscoring the benefits of this decoupling. In our research, we primarily focus on enhancing DRD-type models by incorporating graph information.

The first related area of research is about forecasting daily realized volatility (RV) accounting for spillover effects. Buncic and Gisler (2016); Wang, Pan, and Wu (2018) revealed the cross-market volatility spillover effect from the U.S. equity market to global foreign asset markets, and showed how VIX plays an important role in predicting the volatility of these markets. These volatility spillover effects can be modeled through BEKK, VAR, Wishart autoregression, etc. Nonetheless, Callot, Kock, and Medeiros (2017) pointed out that the aforementioned models may deliver poor out-of-sample forecasts due to the curse of dimensionality. This is mainly because the dependence of each pair of financial assets or markets is modeled as an individual predictor, which leads to substantial computational burdens.

Second, there are some contributions in the existing literature on forecasting daily correlation. In the classic DCC model of Engle (2002), the temporal dependencies in the conditional correlations are characterized by a simple scalar GARCH(1, 1) model, where the dynamics are imposed to be equal for all the correlation pairs. Bauwens and Xu (2023) proposed a novel approach to simultaneous modeling of conditional and realized correlations via DCC-type High-frEquency-bAsed VolatilitY (HEAVY) framework. Generalizations of the above models involve allowing each distinct correlation pair to follow its unique dynamics. For example, Hansen, Lunde, and Voev (2014) introduced a hierarchical structure where the market return is modeled with a univariate realized GARCH model, and individual returns are modeled conditional on the market variables. Bauwens and Otranto (2023) leveraged the element-by-element Hadamard exponential function of the matrix to propose new parameterizations of the conditional autoregressive Wishart (CAW) model family (Bauwens, Storti, and Violante 2012), which imply both asset pair-specific and time-varying impact coefficients of lagged realized correlations via a single parameter.

Despite these well-established studies, it remains somewhat unclear whether and how correlations affect each other, that is, the existence of interdependence of correlation pairs. Within the existing literature, King and Wadhwani (1990) noticed a significant increase in the correlations during crisis periods. Forbes and Rigobon (2002); Aït-Sahalia and Xiu (2016) further revealed that the changes in correlations might be due to a simultaneous increase in the volatility of asset prices.

Third, graph-based models showcased considerable promise in the realm of finance, such as their applications encompassing return prediction (Wang et al. 2022), contagion modeling (Elliott et al. 2014), financial fraud detection (Pourhabibi et al. 2020), etc. Within a financial system, characterized by an array of numerous components with complicated and potentially time-varying relationships, graphs emerge as suitable tools to effectively represent this intricate form of financial data. Acemoglu et al. (2012) argued that idiosyncratic shocks may propagate over supply chains, leading to aggregate fluctuations. Ali and Hirshleifer (2020) identified firm connections by shared analyst coverage, revealing significant momentum spillover effects. Herskovic et al. (2020) documented that network effects are essential to explaining the joint evolution of the empirical firm size and firm volatility distributions, for example, smaller firms and firms with a more concentrated customer base displaying higher volatility. Our article utilizes graph information to capture the interdependence of volatility and correlations.

To some extent, our framework may have some conceptual similarity to the work of Zhu et al. (2017), who proposed a network vector autoregressive model, assuming each node's response at a given time point as a linear combination of (i) its previous value, (ii) the average of its connected neighbors, (iii) a set of node-specific covariates, and (iv) an independent noise. However, our framework addresses the important challenges of constructing desirable graphs, not only for enhancing predictive power on covariances but also for understanding the dynamics of covariances. Our novel framework also allows for time-varying graphs, which can well capture the dynamics in graph effects of volatilities and correlations, while Zhu et al. (2017) assumed a constant graph. Moreover, moving to the covariance matrix forecasting problem opens up a host of interesting and practically relevant econometric applications, as illustrated in our following analysis.

2 Methodologies

This section first outlines the approach for the estimation of RCs, and two traditional models that we use as baselines for forecasting RCs in the U.S. equity market. In Section 2.4, we introduce our new forecasting models based on graphs. Throughout this article, capital bold letters indicate matrices, lowercase bold letters indicate vectors, and plain letters indicate scalars.

2.1 Problem Setup

Let us consider a vector of asset returns $\boldsymbol{r}_t = (r_{1,t}, \ldots, r_{N,t})'$ as follows:

$$\boldsymbol{r}_t = \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \tag{1}$$

where N is the number of assets, and $\boldsymbol{\mu}_t$ is the conditional mean vector given the information set $\mathcal{F}_{t-1}$. $\boldsymbol{\varepsilon}_t$ is a vector of innovations that can be represented as $\boldsymbol{\varepsilon}_t = \Sigma_t^{1/2} \boldsymbol{z}_t$, with $\boldsymbol{z}_t \sim \mathcal{N}(0, \boldsymbol{I}_N)$, where $\boldsymbol{I}_N$ is an $N \times N$ identity matrix. Therefore, Σ_t also represents the conditional covariance matrix of $\boldsymbol{r}_t$.[5]

Note that Σ_t is unobservable. One consistent estimator of Σ_t is the multivariate RC, which is constructed from the intra-day returns, e.g., 5-min returns.[6] Specifically, we denote $\boldsymbol{r}_{t(l)}$ as the l-th vector of non-overlapping Δ-min log-returns during day t, that is, $\boldsymbol{r}_{t(l)} = \log \boldsymbol{p}_{t(l\Delta)} - \log \boldsymbol{p}_{t((l-1)\Delta)}$, where $\boldsymbol{p}_{t(l\Delta)}$ is the price at time $l\Delta$ of day t. Then the RC matrix of day t can be computed as

$$\boldsymbol{H}_t := \sum_{l=1}^{M} \boldsymbol{r}_{t(l)} \boldsymbol{r}'_{t(l)} \tag{2}$$

where $M = 390/\Delta$ is the number of non-overlapping Δ-min slots during day t. In addition, we adopt the subsampling averaging method (see Voev and Lunde 2007; Varneskov and Voev 2013) to improve the above RC estimation, which uses all Δ-minute returns, not just non-overlapping ones.[7]

[5] In this article, our focus is on modeling Σ_t, and we consider $\boldsymbol{\mu}_t$ as determined through separate estimation.

[6] Liu, Patton, and Sheppard (2015) demonstrate that no sub-sampling frequency significantly outperforms a 5-min interval in terms of forecasting daily RVs, making it a widely accepted time interval in the literature.

[7] For example, suppose prices are sampled every minute. The standard realized covariance uses 5-min returns constructed from prices sampled at 09:30, 09:35, 09:40, …, 15:55, 16:00. The subsampled realized covariance uses returns computed from all 5-minute windows, that is, 09:30 and 09:35, 09:31 and 09:36, 09:32 and 09:37, and so on. In the end, the subsampled realized covariance is computed by the average of all standard realized covariances. For other noise-robust estimators of covariances, please refer to Hayashi and Yoshida (2005); Christensen, Kinnebrock, and Podolskij (2010); Barndorff-Nielsen et al. (2011), etc.

The daily realized variance for a particular asset i at day t is $v_{i,t} = \boldsymbol{H}_t[i,i] = \sum_{l=1}^{M} r_{i,t(l)}^2$. We refer to $\boldsymbol{v}_t = (v_{1,t}, \ldots, v_{N,t})'$ as the vector of cross-sectional realized variances. Covariances over longer horizons of k days are computed by the sum of daily RCs (see Symitsi et al. 2018).

2.2 HAR-Cholesky

Corsi (2009) proposed a HAR model for modeling and forecasting the realized variance where the lagged daily, weekly, and monthly variance components are incorporated as predictors. Chiriac and Voev (2011) generalized the univariate HAR model to estimate the RCs as a linear combination of past daily, weekly, and monthly RCs, in the form

$$\boldsymbol{y}_t = \boldsymbol{\alpha}^{(C)} + \beta_d^{(C)} \boldsymbol{y}_{t-1} + \beta_w^{(C)} \boldsymbol{y}_{t-5:t-2} + \beta_m^{(C)} \boldsymbol{y}_{t-22:t-6} + \boldsymbol{u}_t^{(C)}, \tag{3}$$

where $\boldsymbol{y}_t = \mathrm{vech}_{Chol}(\boldsymbol{H}_t)$ denotes the $N^* = N(N+1)/2$ dimensional vectorized version of the Cholesky decomposition of $\boldsymbol{H}_t$ and $\boldsymbol{y}_{t-5:t-2}$ (resp. $\boldsymbol{y}_{t-22:t-6}$) is computed as $\frac{1}{4}\sum_{k=2}^{5} \boldsymbol{y}_{t-k}$ (resp. $\frac{1}{17}\sum_{k=6}^{22} \boldsymbol{y}_{t-k}$), following Symitsi et al. (2018); Bollerslev, Patton, and Quaedvlieg (2018). The intercept $\boldsymbol{\alpha}^{(C)}$ is a N^*-dimensional vector, while the $\beta_d^{(C)}$, $\beta_w^{(C)}$, and $\beta_m^{(C)}$ parameters are all assumed to be scalar. $\boldsymbol{u}_t^{(C)}$ is the error term with conditional mean zeros.

2.3 HAR-DRD

Andersen et al. (2006) demonstrated that the dynamic dependencies in the correlations are different from the volatilities, also known as *correlation breakdown*. Expanding on the seminal work by Engle (2002), Oh and Patton (2016) introduced the HAR-DRD model, which relies on decomposing the RC matrix into the diagonal matrix of realized volatilities and the correlation matrix

$$\boldsymbol{H}_t = \boldsymbol{D}_t \boldsymbol{R}_t \boldsymbol{D}_t, \tag{4}$$

where $\boldsymbol{D}_t$ is the diagonal matrix with the elements of the square roots of $\boldsymbol{v}_t$ on the main diagonal, that is, $\boldsymbol{D}_t[i,i] = \sqrt{v_{i,t}}, \forall i$, and $\boldsymbol{D}_t[i,j] = 0, \forall i \neq j$. $\boldsymbol{R}_t$ is the correlation matrix.

In the HAR-DRD model, the realized variance vector is modeled by vectorized HAR model as in Equation (5).[8] The realized correlation matrix is modeled using a scalar multivariate HAR model as in Equation (6), in line with Bollerslev, Patton, and Quaedvlieg (2018).

$$\boldsymbol{v}_t = \boldsymbol{\alpha}^{(D)} + \beta_d^{(D)} \boldsymbol{v}_{t-1} + \beta_w^{(D)} \boldsymbol{v}_{t-5:t-2} + \beta_m^{(D)} \boldsymbol{v}_{t-22:t-6} + \boldsymbol{u}_t^{(D)}, \tag{5}$$

$$\boldsymbol{x}_t = \boldsymbol{\alpha}^{(R)} + \beta_d^{(R)} \boldsymbol{x}_{t-1} + \beta_w^{(R)} \boldsymbol{x}_{t-5:t-2} + \beta_m^{(R)} \boldsymbol{x}_{t-22:t-6} + \boldsymbol{u}_t^{(R)}, \tag{6}$$

where $\boldsymbol{x}_t = \mathrm{vech}(\boldsymbol{R}_t)$ is the $N^{\#} = N(N-1)/2$ dimensional vectorized version of the lower triangular part of $\boldsymbol{R}_t$ (Henderson and Searle 1979) and $\boldsymbol{x}_{t-5:t-2}$ (resp. $\boldsymbol{x}_{t-22:t-6}$) is computed as $\frac{1}{4}\sum_{k=2}^{5} \boldsymbol{x}_{t-k}$ (resp. $\frac{1}{17}\sum_{k=6}^{22} \boldsymbol{x}_{t-k}$).

[8] An alternative approach involves employing the HAR model for each individual realized volatilities. The outcomes from the individual HARs exhibit slightly inferior performance compared with those derived from Equation (5) (consistent with findings in Bollerslev et al. 2018).

2.4 Enhancing HAR-DRD with Graphs

In the context of financial markets, the price fluctuations of an individual asset are influenced not only by its own characteristics but also by those of similar neighboring assets. However, the HAR-DRD model lacks the ability to capture potential interactions among volatilities and correlations. To address this limitation, we introduce a graph-based approach in which the edges of the graph encode pairwise interactions and inform the model of potential spillover effects, thus ultimately enhancing its predictive power. Before delving into the details of our new forecasting models that incorporate graphs, we lay the groundwork by providing key definitions related to graphs for better clarity and understanding, and to familiarize the reader with these concepts.

2.4.1 Background on graph construction and estimation

Definition 2.1 (Graph). *A graph $\mathcal{G}$ is defined as $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} = \{v_1, \ldots, v_N\}$ is a set of N nodes and $\mathcal{E}$ is a set of edges, where $e_{ij} = (v_i, v_j) \in \mathcal{E}$ denotes an edge connecting node v_i and node v_j.*

Definition 2.2 (Adjacency Matrix). *An adjacency matrix* $\mathbf{A}$ *is a square matrix whose dimension is $N \times N$, where $\mathbf{A}[i,j]$ represents the connection between v_i and v_j in the graph $\mathcal{G}$. An adjacency matrix can be weighted, where $\mathbf{A}[i,j] \geq 0, \forall i,j$ represents the strength/intensity of the connection between nodes v_i and v_j. If $\mathbf{A}[i,j] \in \{0,1\}, \forall i,j$, the graph is a binary graph. In this article, we focus on the**binary** graph.*

Example 1 **(Two Trivial Graphs)**

The adjacency matrix of a **complete** graph (often referred to as K in the literature) contains all ones except along the diagonal where there are only zeros. The adjacency matrix of an **empty** graph is a zero matrix.

Example 2 **(The GICS Sector Graph)**

The Global Industry Classification Standard (GICS) is an industry taxonomy developed for classifying major public companies. A GICS sector graph considers an edge between two asset nodes if they belong to the same GICS sector. Appendix Table A.1 will introduce the corresponding sector of each stock under consideration.

GL

GL is a sparsity-penalized maximum likelihood estimator for the precision matrix (i.e., inverse of the covariance matrix) proposed by Friedman, Hastie, and Tibshirani (2008). Assume observations of $\boldsymbol{\varepsilon}_t$ in Equation (1) are drawn from a multivariate Gaussian distribution $\mathcal{N}(0, \Sigma)$. If the ij-th component of the precision matrix is zero, then the i-th and j-th variables are conditionally independent, given the other variables (for further details, see Meinshausen and Bühlmann 2006; Friedman, Hastie, and Tibshirani 2008). Alternatively, zeros in the precision matrix indicate conditional independence. Thus we are interested in estimating the precision matrix $\mathbf{\Theta} = \mathbf{\Sigma}^{-1}$. The GL algorithm aims to obtain an estimator $\hat{\Theta}$ that minimizes the following objective function:

$$\hat{\Theta} = \mathrm{argmin}_{\Theta \succeq 0}\Big(\mathrm{Tr}(\boldsymbol{S}\Theta) - \log\det(\Theta) + \lambda \sum_{j \neq k} |\Theta_{jk}|\Big),$$

where $\boldsymbol{S}$ is the sample covariance of $\boldsymbol{\varepsilon}_t$, and λ is the penalizing trade-off parameter that controls the sparsity of $\hat{\Theta}$. In this work, we use the standard five-fold cross-validation procedure to choose λ based on training data. We remove these zero edges in the estimated precision matrix (i.e., those connecting conditionally independent node pairs) and consider the non-zero elements as edges. In other words, the adjacency matrix $\boldsymbol{A}$ from GL is defined as $\boldsymbol{A}[i,j] = 1$ if $\hat{\Theta}[i,j] \neq 0$; otherwise $\boldsymbol{A}[i,j] = 0$.

Definition 3.3 (Line Graph). *Given a graph $\mathcal{G}$, its line graph $L(\mathcal{G})$ is a graph such that*

- *each node of $L(\mathcal{G})$ represents an edge of $\mathcal{G}$;*
- *two nodes of $L(\mathcal{G})$ are adjacent if and only if their corresponding edges share a common endpoint in $\mathcal{G}$.*

Figure 1 shows an illustration of graphs defined in our framework. Figure 1a models the interdependence between the volatility of assets, where each node represents an asset. Their edges, as shown in Figure 1b, constitute a node in the line graph shown in Figure 1c. The edges in the line graph capture the interdependence between two correlation pairs that have an asset in common.

Based on the HAR-DRD model in Equation (4), we propose to forecast the RC matrix by modeling the realized volatilities and correlation matrix separately. For each sub-task, we incorporate new predictors that represent the structural information contained in the aforementioned graphs via neighborhood aggregation.

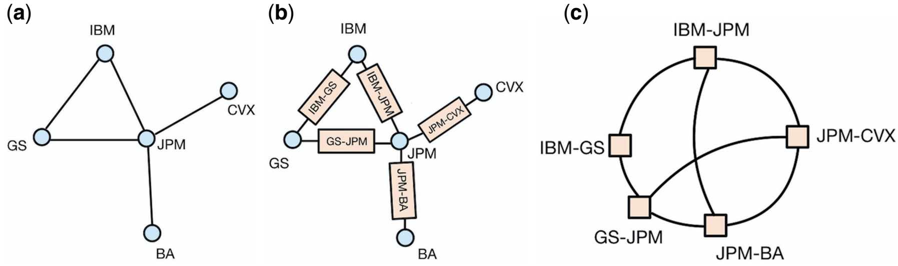
2.4.2 Forecasting realized volatilities with graphs

The HAR model assumes that the future volatility of a specific asset only relies on its own past volatilities. Considering that volatilities of connected assets may have predictive power on the volatility of the target asset, we propose the following model to aggregate the spillover effects from neighbors (i.e., connected assets), which we denote as the Graph-HAR model (or in short GHAR)

$$\begin{aligned}\mathbf{GHAR}(\boldsymbol{A}) : \boldsymbol{v}_t = \boldsymbol{\alpha}^{(D)} &+ \underbrace{\beta_d^{(D)}\boldsymbol{v}_{t-1} + \beta_w^{(D)}\boldsymbol{v}_{t-5:t-2} + \beta_m^{(D)}\boldsymbol{v}_{t-22:t-6}}_{Self} \\ &+ \underbrace{\gamma_d^{(D)}\boldsymbol{W}\cdot\boldsymbol{v}_{t-1} + \gamma_w^{(D)}\boldsymbol{W}\cdot\boldsymbol{v}_{t-5:t-2} + \gamma_m^{(D)}\boldsymbol{W}\cdot\boldsymbol{v}_{t-22:t-6}}_{Graph} + \boldsymbol{u}_t^{(D)},\end{aligned} \tag{7}$$

where $\boldsymbol{W} = \boldsymbol{O}^{-\frac{1}{2}}\boldsymbol{A}\boldsymbol{O}^{-\frac{1}{2}}$ is the normalized adjacency matrix.[9] Specifically, $\boldsymbol{A}$ is a $N \times N$ adjacency matrix indicating the connections between assets with diagonal elements as 0, and $\boldsymbol{O} = \mathrm{diag}\{n_1, \ldots, n_N\}$, where $n_i = \sum_j \boldsymbol{A}[i,j], \forall i$. Therefore $\boldsymbol{W}\cdot\boldsymbol{v}_{t-1}, \boldsymbol{W}\cdot\boldsymbol{v}_{t-5:t-2}$, $\boldsymbol{W}\cdot\boldsymbol{v}_{t-22:t-6}$ represent the **neighborhood aggregation** over daily, weekly, and monthly horizons. γ_d, γ_w, γ_m represent the effects from connected neighbors over different horizons.

[9] We also conducted experiments without normalization, and the results were quite similar. However, normalization facilitates a straightforward comparison between the effects of *Self* and *Graph*, thereby enhancing the interpretability of the model.



Graph $\mathcal{G}$ for volatility Labeled edges in $\mathcal{G}$ Line graph $L(\mathcal{G})$ for correlation

Figure 1 An illustration of the process of building the line graph $L(\mathcal{G})$ for $N=5$ assets.

For simplicity, we denote the effect from a stock's own historical events as *Self*, and that from a graph as *Graph*, as shown in Equation (7). We use the Ordinary Least Squares (OLS) to obtain the estimates of Equation (7).

In terms of the choice of adjacency matrices for variance modeling, we first consider an empty graph, that is, the elements of $\boldsymbol{A}$ are all 0 s. In this case, Equation (7) reduces to the standard vectorized HAR model (5) for modeling volatilities. When the off-diagonal elements of $\boldsymbol{A}$ are all 1 s, that is, a complete graph, $\boldsymbol{W}\cdot\boldsymbol{v}_{t-1}$ represents the global variance as studied in Bollerslev et al. (2018). Given that companies from the same sector tend to behave similarly and also influence each other, the GICS sector is another frequently used knowledge-based graph for the construction of $\boldsymbol{A}$, e.g., Fan, Furger, and Xiu (2016). In addition, we employ a data-driven approach, that is, GL, to compute the adjacency matrix (see Hallac et al. 2017).

2.4.3 Forecasting realized correlations with graphs

Moreover, we apply the idea of the graph effect to modeling correlations according to the model

$$
\begin{aligned}
\textbf{GHAR}(\widetilde{\boldsymbol{A}}):\boldsymbol{x}_t = \boldsymbol{\alpha}^{(R)} + &\underbrace{\beta_d^{(R)}\boldsymbol{x}_{t-1} + \beta_w^{(R)}\boldsymbol{x}_{t-5:t-2} + \beta_m^{(R)}\boldsymbol{x}_{t-22:t-6}}_{Self} \\
&+ \underbrace{\gamma_d^{(R)}\widetilde{\boldsymbol{W}}\boldsymbol{x}_{t-1} + \gamma_w^{(R)}\widetilde{\boldsymbol{W}}\boldsymbol{x}_{t-5:t-2} + \gamma_m^{(R)}\widetilde{\boldsymbol{W}}\boldsymbol{x}_{t-22:t-6}}_{Graph} + \boldsymbol{u}_t^{(R)},
\end{aligned}
\tag{8}
$$

where $\widetilde{\boldsymbol{W}} = \widetilde{\boldsymbol{O}}^{-\frac{1}{2}}\widetilde{\boldsymbol{A}}\ \widetilde{\boldsymbol{O}}^{-\frac{1}{2}}$ is the normalized adjacency matrix. Specifically, $\widetilde{\boldsymbol{A}}$ is a $N^{\#}\times N^{\#}$ ($N^{\#}=N(N-1)/2$) adjacency matrix indicating the connections between pairwise correlations with diagonal elements as 0, and $\widetilde{\boldsymbol{O}}=\mathrm{diag}\{\widetilde{n}_1,\ldots,\widetilde{n}_{N^{\#}}\}$, where $\widetilde{n}_i=\sum_j\widetilde{\boldsymbol{A}}[i,j], \forall i$.

In terms of adjacency matrices of correlation networks, in addition to the empty graph (all 0's) and complete graph (all 1's), we also adopt the line graph, where the correlation $\boldsymbol{R}[i,j]$ between asset i and j ($i\neq j$) is connected to another correlation $\boldsymbol{R}[k,l]$ between k and l ($k\neq l$), iff $\{i,j\}\cap\{k,l\}\neq\emptyset$ and $\{i,j\}\neq\{k,l\}$. It is not difficult to obtain that the ratio between the degree of a line graph and the degree of a complete graph is $\frac{N^{\#}(N-2)}{N^{\#}(N^{\#}-1)/2}=\frac{4(N-2)}{N(N-1)-2}$. To the best of our knowledge, this is the first study to explore the predictive information of line graphs for forecasting realized correlation matrices.

As a notational convention, we use $\boldsymbol{A}$ (resp. $\widetilde{\boldsymbol{A}}$) to denote the adjacency matrix for volatility (resp. correlation). Additionally, we refer to GHAR($\boldsymbol{A}$) (resp. GHAR($\widetilde{\boldsymbol{A}}$)) as the augmented HAR model with graph information $\boldsymbol{A}$ (resp. $\widetilde{\boldsymbol{A}}$) for forecasting volatilities

Table 1. Choices of adjacency matrices for modeling volatilities and correlations, and the joint models for forecasting RCs

$A \backslash \widetilde{A}$	Empty	Complete	Line
Empty	HAR-DRD	GHAR(-, $\widetilde{K}$)	GHAR(-, $\widetilde{L}$)
Complete	GHAR(K, -)	GHAR(K, $\widetilde{K}$)	GHAR(K, $\widetilde{L}$)
Sector	GHAR(S, -)	GHAR(S, $\widetilde{K}$)	GHAR(S, $\widetilde{L}$)
GLASSO	GHAR(GL, -)	GHAR(GL, $\widetilde{K}$)	GHAR(GL, $\widetilde{L}$)

(resp. correlations). In the end, GHAR($\boldsymbol{A}$,$\widetilde{\boldsymbol{A}}$) is referred to as the model for forecasting RCs, which combines the forecasts of realized volatilities based on $\boldsymbol{A}$ and correlations based on $\widetilde{\boldsymbol{A}}$. Table 1 lists the adjacency matrices for realized volatilities and correlations, and the corresponding joint models.

3 Empirical Analysis

We obtain intraday data on the components of Dow Jones Industrial Average (DJIA) from LOBSTER for the period from July 1, 2007, to July 1, 2021.[10] Following Bollerslev, Patton, and Quaedvlieg (2016), only stocks that traded continuously from the start to the end of our sample are maintained. As a result, twenty-seven Dow Jones constituents are in the final sample and their ticker symbols are summarized in Appendix A, which also provides relevant statistics for the volatility/correlation estimates.

3.1 Forecast Evaluation Metrics

Following the convention in the literature, we use hats to denote forecasts. For example, we refer to the forecast of the conditional covariance matrix on day t as $\hat{\Sigma}_t$. Consistent with previous studies (e.g., Symitsi et al. 2018; Bollerslev, Patton, and Quaedvlieg 2018), we consider the following loss functions to measure the average distance between predicted covariances and RCs for the model comparisons.[11]

$$\mathcal{L}_t^E = \sqrt{\mathrm{vech}(\boldsymbol{H}_t - \hat{\Sigma}_t)' \mathrm{vech}(\boldsymbol{H}_t - \hat{\Sigma}_t)}, \tag{9}$$

$$\mathcal{L}_t^F = \sqrt{\mathrm{Tr}[(\boldsymbol{H}_t - \hat{\Sigma}_t)'(\boldsymbol{H}_t - \hat{\Sigma}_t)]},$$
$$\mathcal{L}_t^Q = \log\det(\hat{\Sigma}_t) + \mathrm{Tr}[\hat{\Sigma}_t^{-1}\boldsymbol{H}_t].$$

Here $\mathcal{L}_t^E$ represents the Euclidean distance between the vectorization of forecast covariances and ex-post RCs. $\mathcal{L}_t^F$ is the Frobenius distance between the two matrices. $\mathcal{L}_t^Q$ is the quasi-likelihood (QLIKE) loss based on the negative log-likelihood of a multivariate normal. All of these functions measure losses; therefore, *lower values are preferred.*

[10] LOBSTER database contains data from the June 27, 2007, up to the day before yesterday. LOBSTER is subject to licensing restrictions and must be purchased by users, as redistribution is prohibited.

[11] The below equation illustrates the inherent relationship between the Euclidean and Frobenius distance (also see Bauwens and Otranto (2023)). Generally, the conclusions derived from $\mathcal{L}_t^E$ and $\mathcal{L}_t^F$ should be similar. However, it is important to note that $\mathcal{L}_t^E$ places more emphasis on the diagonal elements, representing volatilities, compared to $\mathcal{L}_t^F$. Regarding the relation/difference between $\mathcal{L}_t^E$ ($\mathcal{L}_t^F$) and $\mathcal{L}_t^Q$, $\mathcal{L}_t^E$ and $\mathcal{L}_t^F$ depend on the forecast error $\boldsymbol{H}_t - \hat{\Sigma}_t$, while $\mathcal{L}_t^Q$ depends on the ratio $\hat{\Sigma}_t^{-1}\boldsymbol{H}_t$. $\mathcal{L}_t^E = \sqrt{[\sum_i (H_{ii,t} - \hat{\Sigma}_{ii,t})^2 + \sum_{i<j}(H_{ij,t} - \hat{\Sigma}_{ij,t})^2]}$, $\mathcal{L}_t^F = \sqrt{[\sum_i (H_{ii,t} - \hat{\Sigma}_{ii,t})^2 + 2\sum_{i<j}(H_{ij,t} - \hat{\Sigma}_{ij,t})^2]}$.

3.1.1 MCS

Hansen, Lunde, and Nason (2011) proposed a nonparametric statistical test to formally compare the forecasting accuracy of different models, denoted as MCS. The MCS identifies a set of models that contains the best forecasting model(s) with a given level of confidence. Specifically, given a collection of models $\mathcal{M}_0$, the test sequentially discards models with inferior predictive ability, and in the end, determines a subset containing the best model(s) $\mathcal{M}^*$. The iterative elimination is based on sequentially testing the following hypothesis

$$H_0 : \mathbb{E}(\Delta\mathcal{L}_{ij,t}) = 0, \text{for all } i, j \in \mathcal{M}^*, \tag{10}$$

where $\mathcal{L}_{ij,t}$ is the loss difference between models i and j at day t in terms of a specific loss function $\mathcal{L}$ (from Equation (9)). The MCS procedure makes it possible to make statements about statistical significance from multiple pairwise comparisons. For additional details, we refer to the studies of Hansen, Lunde, and Nason (2003, 2011). Consistent with Symitsi et al. (2018), we implement the MCS procedure using a block bootstrap procedure with 10,000 replications and a block length of 2 observations.

3.2 In-Sample Results

Following the workflow in Bollerslev, Patton, and Quaedvlieg (2016, 2018), we start the empirical analysis with an in-sample test of each model under consideration, followed by a graph structure study.

3.2.1 Model evaluation

Table 2 reports the estimated coefficients for modeling the 1-day-ahead RCs under the Cholesky decomposition, and the estimated coefficients for modeling 1-day-ahead realized volatilities and correlations under the DRD decomposition. The HAR-Cholesky model attributes the most importance to weekly RVs, followed by the monthly and daily components.[12] Additionally, our results support a stronger effect of the first lag on volatilities ($\beta_d = 0.609$) than the respective one on correlations ($\beta_d = 0.289$), consistent with the existing literature (e.g., Bollerslev, Patton, and Quaedvlieg 2018).

When examining the graph effect, we find the GHAR model for forecasting volatilities redistributes some weights from a stock's own daily lag to the daily lag of its connected neighbors. We observe a similar shift for correlation forecasting. Interestingly, the estimated coefficients of the weekly and monthly lags from neighbors are negative and statistically significant in the in-sample test. Future studies on the negative graph effects of longer lags are therefore recommended.

Table 3 reports the average in-sample predictive performance of each model under consideration. To conserve space, we present (i) the ratio of loss functions when comparing each model with HAR-DRD; (ii) the rank of loss values in thirteen model candidates; (iii) p_{MCS} of the MCS test. In this manner, a value of ratio less than 1 indicates the predictive performance is better than the baseline model HAR-DRD; a rank of 1 indicates the model has the best in-sample fit; p-val greater than 5% indicates the model(s) that are superior to all other models identified by MCS.

From Table 3, we summarize the following findings. First, we conclude that HAR-Cholesky produces worse results than HAR-DRD, which is consistent with the conclusion of Bollerslev, Patton, and Quaedvlieg (2018). The results further show that the GHAR(GL, $\widetilde{L}$) model achieves the best in-sample fit across all loss functions. For example, looking at the value of the $\mathcal{L}^E$ (resp. $\mathcal{L}^F$, $\mathcal{L}^Q$) loss function, we observe that the average loss of GHAR (GL, $\widetilde{L}$) is about 98.3% (resp. 98.4%, 98.2%) of the baseline, HAR-DRD. In general, all

[12] The HAR-Cholesky in Bollerslev, Patton, and Quaedvlieg (2018) model allocates roughly 0.247, 0.410, and 0.244 to the daily, weekly, and monthly components, respectively.

Table 2. In-sample estimation

		β_d	β_w	β_m	γ_d	γ_w	γ_m
Covariance	HAR-Cholesky	0.311	0.413	0.242			
		(0.001)	(0.001)	(0.001)			
Volatility	HAR	0.609	0.271	0.057			
		(0.003)	(0.004)	(0.003)			
	GHAR(K)	0.483	0.279	0.169	0.239	−0.075	−0.150
		(0.004)	(0.006)	(0.005)	(0.006)	(0.008)	(0.006)
	GHAR(S)	0.492	0.277	0.155	0.187	−0.048	−0.120
		(0.004)	(0.006)	(0.006)	(0.005)	(0.007)	(0.006)
	GHAR(GL)	0.471	0.285	0.144	0.224	−0.075	−0.103
		(0.004)	(0.006)	(0.006)	(0.005)	(0.007)	(0.006)
Correlation	HAR	0.289	0.386	0.213			
		(0.001)	(0.001)	(0.001)			
	GHAR($\widetilde{K}$)	0.144	0.268	0.484	0.348	0.068	−0.413
		(0.001)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)
	GHAR($\widetilde{L}$)	0.076	0.222	0.595	0.407	0.121	−0.521
		(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.004)

Note: The table reports the in-sample estimates for coefficients and their standard errors (in parentheses) of every model, calculated from the entire sample period. The Covariance row reports the HAR-Cholesky model, for forecasting the Cholesky decomposition vector of the RC matrix. The Volatility rows include different models for forecasting multivariate realized variances, while the Correlation rows include different models for forecasting the lower triangular part of the realized correlation matrix.

Table 3. In-sample losses.

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	Rank	p_{MCS}	Ratio	Rank	p_{MCS}	Ratio	Rank	p_{MCS}
HAR-Cholesky	**1.023**	13	<0.001	1.032	13	<0.001	1.051	13	<0.001
HAR-DRD	**1.000**	12	<0.001	1.000	12	<0.001	1.000	10	<0.001
GHAR(-, $\widetilde{K}$)	**0.997**	11	<0.001	0.997	11	<0.001	0.987	6	<0.001
GHAR(-, $\widetilde{L}$)	**0.995**	10	<0.001	0.994	10	<0.001	0.985	4	<0.001
GHAR(S, -)	**0.990**	9	<0.001	0.992	9	<0.001	1.001	11	<0.001
GHAR(S, $\widetilde{K}$)	**0.988**	8	<0.001	0.989	8	<0.001	0.987	5	<0.001
GHAR(S, $\widetilde{L}$)	**0.986**	7	<0.001	0.987	5	<0.001	0.985	3	<0.001
GHAR(K, -)	**0.986**	5	<0.001	0.987	6	<0.001	1.006	12	<0.001
GHAR(K, $\widetilde{K}$)	**0.985**	4	<0.001	0.986	4	<0.001	0.991	8	<0.001
GHAR(K, $\widetilde{L}$)	**0.983**	2	0.557	0.984	2	0.745	0.989	7	<0.001
GHAR(GL, -)	**0.986**	6	<0.001	0.988	7	<0.001	0.998	9	<0.001
GHAR(GL, $\widetilde{K}$)	**0.984**	3	<0.001	0.986	3	<0.001	0.985	2	<0.001
GHAR(GL, $\widetilde{L}$)	**0.983**	1	1.000	0.984	1	1.000	0.982	1	1.000

Note: The table reports the in-sample losses of various models over the 1-day forecast horizons, averaged over the entire sample. $\mathcal{L}^E$ is the Euclidean distance, $\mathcal{L}^F$ is the Frobenius distance, and $\mathcal{L}^Q$ is the Quasi-Likelihood loss function.

graph models outperform the baseline, providing strong evidence for the importance of leveraging graph information and the existence of certain cross-asset interaction effects.

3.2.2 Graph analysis

Figure 2 displays the adjacency matrices of various graphs for volatilities. The GICS sector graph in Figure 2b has a block-diagonal structure, with fewer edges than the GL and

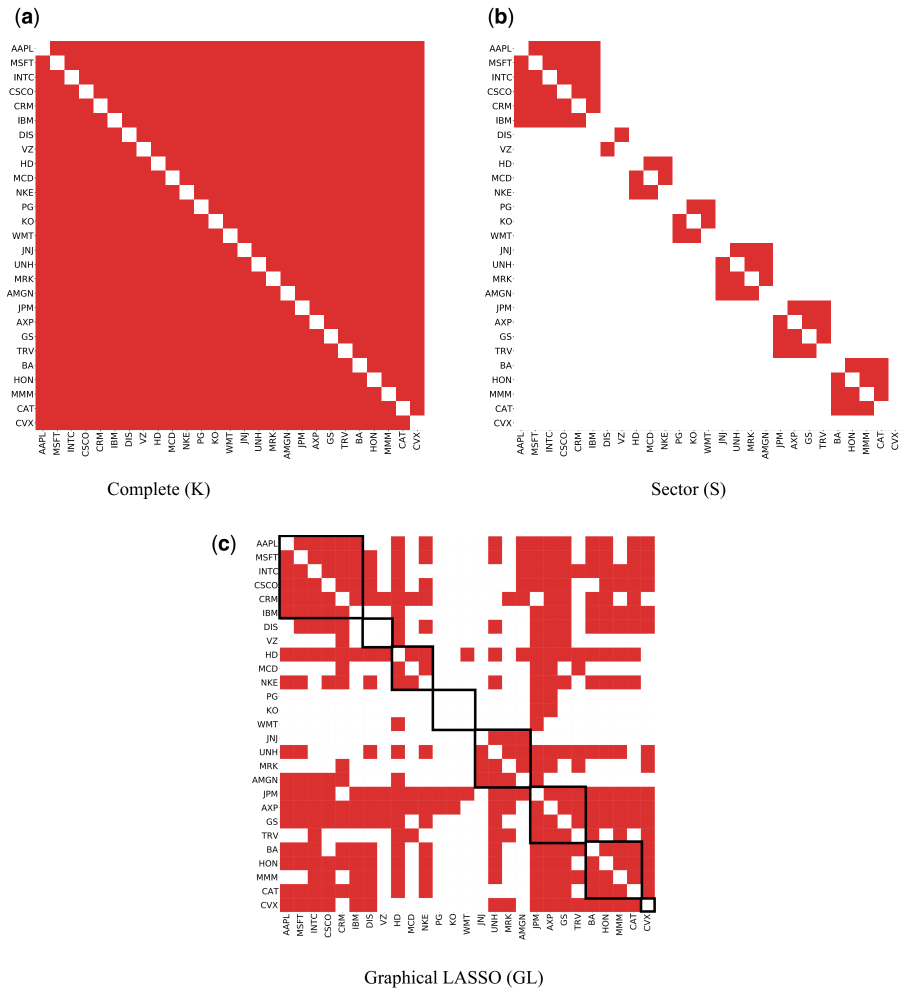
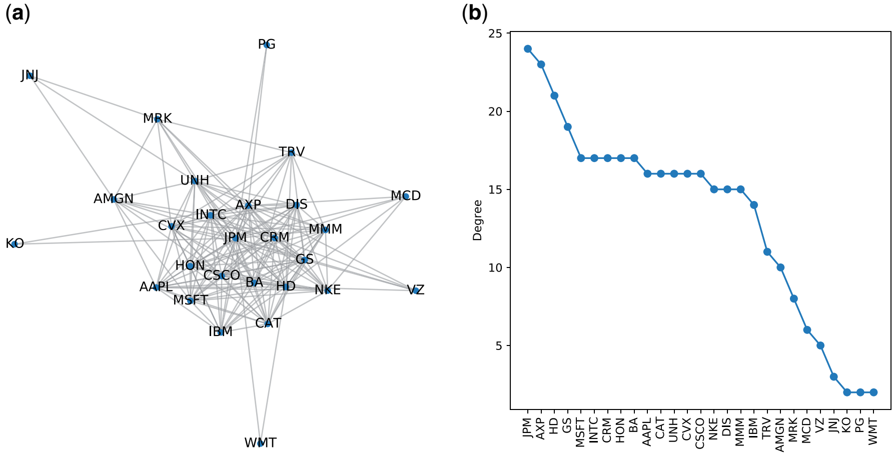


Figure 2 Adjacency matrices for realized variances. The figure shows the adjacency matrices from various methods, for forecasting realized variances. The stocks are sorted by their industrial sectors, as listed in Appendix Table A.1. A value of 1 represents a connection between two stocks, while a value of 0 indicates no connection.

complete graphs. The neighborhood aggregation via the GICS sector graph in Equation (7) is equivalent to calculating the average volatility within each sector, while ignoring the inter-sector links.

The adjacency matrix chosen by GL (shown in Figure 2c) maintains a GICS sector structure to some extent, such as information technology, consumer discretionary, health care, financials, and industrials. Meanwhile, we observe that there are substantial inter-sector connections, especially in financials and information technology, to capture market information.

To better illustrate the interactions between stocks modeled by GL, we construct a network based on the adjacency matrix of 2(c), shown in Figure 3a. We also calculate the node degrees and sort stocks in descending order of degree. Figure 3b reveals that financial stocks, such as JPM, AXP, and GS, appear to be the center of the volatility network,



Connected components

Degree-rank plot

Figure 3 Illustration of the graph corresponding to the adjacency matrix in 2(c) and degree-rank plot.

followed by technology companies, including MSFT, INTC, etc. This observation could provide insight into the superior in-sample fitness of the GL approach compared to the GICS sector-based approach. Fixed sector weights rely on predefined sector classifications, which may not accurately capture the underlying relationships among assets. The GL approach avoids making strong assumptions about sector-based correlations and instead learns relationships directly from data. For example, from Figure 3b, nearly every asset is connected to JPM, one of the largest banks. This connectivity can be attributed to the recognition of JPM's systemic significance within the broader financial market, a crucial aspect that remains unaccounted for within the purview of the GICS sector approach.

3.3 Out-of-Sample Results

We now focus our discussion on the out-of-sample predictive performance of the competing models in a rolling window setup.

3.3.1 Model evaluation

Following Bollerslev, Patton, and Quaedvlieg (2016, 2018), Symitsi et al. (2018), and Pascalau and Poirier (2021), we estimate the parameters of each model using a fixed length rolling window of the most recent 1000 observations and obtain the rolling 1-day forecasts for each trading day in the next month. To this end, the out-of-sample period lasts from July 2011 to July 2021. All of the models are re-estimated every month.[13]

To ensure that the predicted correlations lie in $[-1, 1]$, we winsorize the values, meaning that we set the values below -1 to -1, and those above 1 to 1 in the following experiments.[14] Furthermore, to ensure the positive definiteness of covariances, we find the nearest positive definite covariance matrix to replace the negative definite covariance

[13] In Appendix D, we also report the performance of the models updated more frequently, such as daily and weekly, which leads to similar conclusions.

[14] Another common method is to use some transformation functions to convert volatilities and correlations before forecasting, e.g., log transformation for volatilities (see Andersen et al. (2003); Bucci (2020a)) and Fisher transformation for correlations (see Andersen et al. (2006)). We report the results in Appendix E.

Table 4. Out-of-sample losses over 1-day forecast horizon

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}
HAR-Cholesky	1.031	13	< 0.001	1.032	13	< 0.001	1.071	13	< 0.001
HAR-DRD	1.000	12	0.002	1.000	12	0.003	1.000	9	< 0.001
GHAR(-, $\widetilde{K}$)	0.993	11	0.002	0.991	11	0.003	0.986	7	< 0.001
GHAR(-, $\widetilde{L}$)	0.991	10	0.004	0.989	9	0.006	0.984	3	0.048
GHAR(S, -)	0.990	9	0.002	0.990	10	0.003	1.002	12	< 0.001
GHAR(S, $\widetilde{K}$)	0.983	8	0.003	0.983	7	0.005	0.987	8	< 0.001
GHAR(S, $\widetilde{L}$)	0.982	7	0.004	0.981	5	0.006	0.985	6	< 0.001
GHAR(K, -)	0.982	5	0.004	0.983	6	0.006	1.001	11	< 0.001
GHAR(K, $\widetilde{K}$)	0.976	4	0.004	0.976	3	0.006	0.985	5	< 0.001
GHAR(K, $\widetilde{L}$)	0.975	2	0.869	0.975	1	1.000	0.982	2	0.277
GHAR(GL, -)	0.982	6	0.004	0.983	8	0.006	1.001	10	< 0.001
GHAR(GL, $\widetilde{K}$)	0.976	3	0.005	0.976	4	0.006	0.984	4	< 0.001
GHAR(GL, $\widetilde{L}$)	0.975	1	1.000	0.975	2	0.768	0.982	1	1.000

Note: The table reports the out-of-sample losses of various models over the 1-day forecast horizon, averaged over the entire testing sample. $\mathcal{L}^E$ is the Euclidean distance, $\mathcal{L}^F$ is the Frobenius distance, and $\mathcal{L}^Q$ is the Quasi-Likelihood loss function.

matrices.[15] We utilize the cov_nearest function in the built-in Python package statsmodels[16] to compute the nearest covariance matrix, which converts the covariance matrix to a correlation matrix, then finds the nearest positive definite correlation matrix, and converts it back to a covariance matrix using the volatility forecasts.[17] This function iteratively adjusts the correlation matrix by clipping the eigenvalues of an intermediate matrix such that the smallest eigenvalue is approximately equal to the predetermined threshold.

Table 4 illustrates the average losses over the entire out-of-sample period, that is, July 2011–July 2021. We first focus on the comparison between HAR-Cholesky and HAR-DRD. Consistent with the in-sample results, HAR-DRD significantly outperforms HAR-Cholesky across all losses. Furthermore, the results show that the GHAR(GL, $\widetilde{L}$) model yields the most accurate out-of-sample forecasts across all loss functions. Specifically, based on the $\mathcal{L}_E$ (resp. $\mathcal{L}_F$, $\mathcal{L}_Q$) loss function, the GHAR(GL, $\widetilde{L}$) model has about 2.5% (resp. 2.5%, 1.8%) lower average forecast error compared to the baseline HAR-DRD. To disentangle the prediction improvements in covariances, we compare the performance of HAR-DRD, GHAR(GL, -), GHAR(-, $\widetilde{L}$), and GHAR(GL, $\widetilde{L}$). We observe that the graph information helps improve the predictive performance in both volatilities and correlations forecasting, which subsequently leads to an improved prediction of the covariance matrix. Indeed, GHAR(GL, $\widetilde{L}$), which utilizes both graphs in volatility and correlation simultaneously, clearly yields superior results compared to the models that only utilize one separate graph either in volatility or correlation. The MCS test indicates that in addition to GHAR(GL, $\widetilde{L}$), the GHAR(K, $\widetilde{L}$) model is also included in the subset of best models, across all loss functions. The superiority of these two models may be attributed to the high level

15 We also investigated the insanity filter method proposed in Bollerslev, Patton, and Quaedvlieg (2018), which replaces the negative definite covariance matrices with the simple average realized covariance matrix over the relevant estimation sample period.

16 https://www.statsmodels.org

17 The insanity filter method of Bollerslev, Patton, and Quaedvlieg (2018) might oversimplify the data structure and lose temporal dependencies, especially with trending data. In contrast, the cov_nearest method transforms the correlation matrix while keeping the variances unchanged. This approach also preserves a greater degree of the original correlation structure by clipping small eigenvalues to a predetermined threshold. Moreover, it can better capture the temporal dependencies, since it relies on the raw forecasts from the HAR model. Nonetheless, the empirical performance of these two methods is similar, given the rarity of negative definite covariance matrices occurring throughout the entire sample for each GHAR model variant.

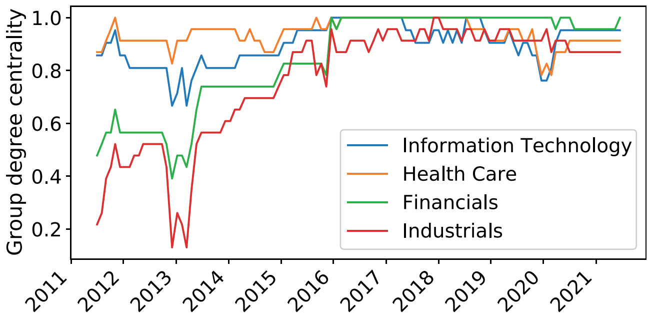


Figure 4 Group degree centrality of four sectors in the graphs from GL.

of interconnectedness among assets' volatilities within our studied universe, as captured by the GL approach in Figures 2 and 3. Additionally, the effective incorporation of the line graph for correlations further contributes to their competitive edge.

In addition, GHAR(S, ·)[18] has an inferior predictive performance, compared to GHAR (GL, ·). Beyond the potential limitations in accurately capturing underlying relationships posed by the sector approach, an additional explanation might be that the adaptive and time-varying GL approach allows the modeling of **dynamic relationships** among assets over time, that might be missed by fixed sector weights. Thus, the GL approach can also potentially capture **emerging trends** and relationships that fixed sector weights might overlook (refer to Figure 4). Furthermore, financial markets are prone to **structural changes** due to various factors, such as economic events or policy shifts. The GL approach can adjust the graph structures to account for these changes, as shown in Figure 5. This also motivates us to examine the dynamics of graphs captured by GL in the following subsection.

We further investigate whether the coefficients in each model vary across time. Figure 6 displays the rolling 4-year betas (and gammas) and their respective 95% confidence intervals for modeling volatilities over the testing period. The plots of HAR first show that there are three substantial shocks in β_d, which occur at the end of 2012, the end of 2015, and March 2020. In all GHAR models, β_d rapidly drops at the end of 2012 and 2015 as well. However, β_d appears to be more consistent during the pandemic period (March 2020), while the graph coefficient γ_d displays a high jump during the COVID-19 pandemic. This may indicate that the spillover effect is a major driver of market volatility during the pandemic. Another surprising aspect of the plots is that β_w has a similar trend across various models, and similarly for β_m. We also observe that γ_w is insignificant during 2016–2020 at the 5% confidence level. These findings indicate that lagged weekly and monthly graph effects may have a minor influence on volatility dynamics.[19]

We provide a similar analysis of the models for correlations, illustrated in Figure 7. First, the rolling coefficients appear to be less volatile than their counterparts when modeling volatilities. As expected, the rolling coefficients look similar across graph-based methods. In addition, we observe that the coefficients β_d, β_w, and β_m are predominantly smoother for graph-based models compared to their counterparts in the standard HAR. Interestingly, the lagged daily graph effect γ_d in graph-based models exhibits a similar trend as β_d in HAR.

3.3.2 Graph analysis

Since GHAR(GL, $\widetilde{L}$) delivers the best forecasting results, we further dive into the dynamics of the graphs uncovered by GL. For each adjacency matrix at a timestamp, we calculate the

18 We use the up-to-date GICS sector to construct a static graph employed during the entire period.

19 We study the graph-based model with only lagged daily graph effect in Appendix C.

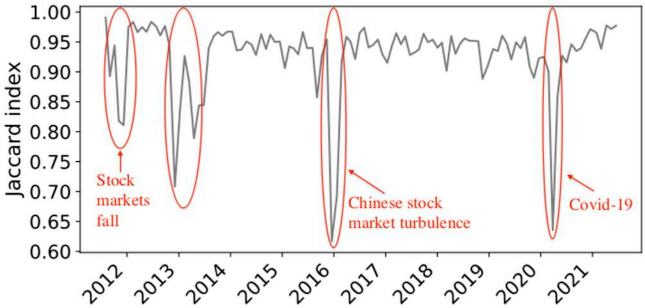


Figure 5 Jaccard index of similarity between consecutive graphs.

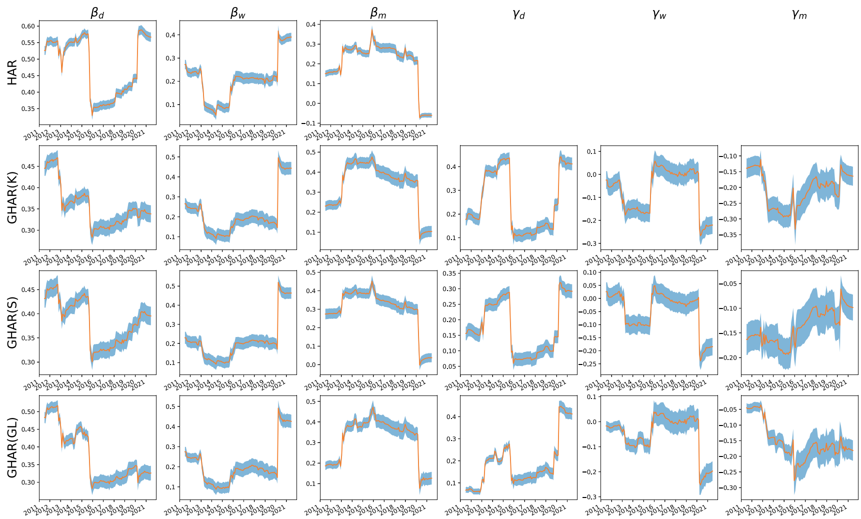


Figure 6 Rolling coefficients of various models for forecasting volatilities. The central curve represents the estimated coefficients, with the shaded bands covering the 95% confidence interval.

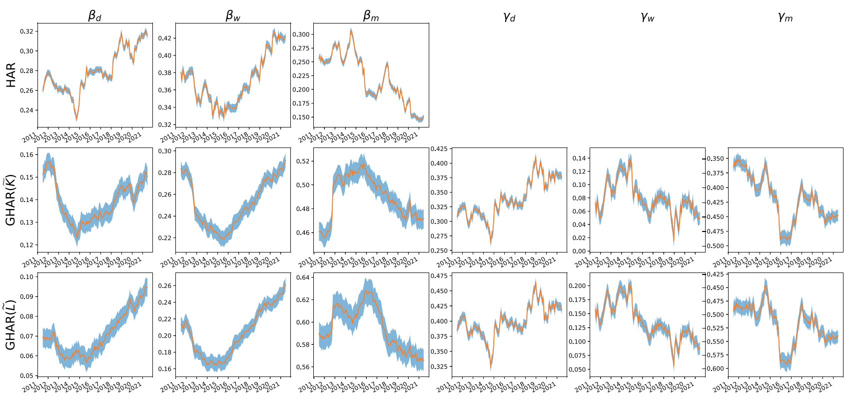


Figure 7 Rolling coefficients of various models for forecasting correlations. The central curve represents the estimated coefficients, with the shaded bands covering the 95% confidence interval.

fraction of non-sector stocks that are connected to sector members, a.k.a *group degree centrality* (see more in Everett and Borgatti 1999). For ease of visualization, in Figure 4, we only plot group degree centrality of four GICS sectors with most constituents. We first observe that the centrality for each sector maintains at a higher level since 2015, implying that sector connectedness becomes increasingly important in the U.S. stock market, in line with the findings of Baruník, Kočenda, and Vácha (2016) and Costa, Matos, and da Silva (2022). In addition, we observe that at the beginning of the COVID-19 pandemic, healthcare and information technology companies exhibit increasing influence on other stocks, which is not surprising as they are two critical sectors whose prices have been largely affected during the pandemic.

To study the stability and temporal dynamics of graphs learned by GL, we also calculate the similarity of two consecutive graphs at monthly basis in terms of their edge sets. The Jaccard index of the edge set is defined as the ratio between the number of intersection edges and the number of union edges of two consecutive graphs. Specifically, at time t, the edge set of the GL estimate is denoted as $\mathcal{E}^{(t)}$, with the Jaccard index computed as $\frac{|\mathcal{E}^{(t)} \cap \mathcal{E}^{(t-1)}|}{|\mathcal{E}^{(t)} \cup \mathcal{E}^{(t-1)}|}$.

We plot the Jaccard index between consecutive GL estimates against the timespan in Figure 5. The graphs are stable in general, with only a few exceptions: Aug 2011–Dec 2011, Nov 2012–June 2013, Dec 2015–Feb 2016, and March 2020–April 2020. This suggests that the equity market might be experiencing significant changes during these periods. In fact, some periods correspond to stock market turmoil, as highlighted in Figure 5. To the best of our knowledge, there was no major financial event from the end of 2012 until mid-2013 documented in the literature, yet very recent methods in the change-point detection machine learning literature have also picked up this regime (see Sulem et al. 2024).

3.4 Portfolio Performance

To further quantify the benefit of our proposed GHAR models, we examine the portfolio performance based on the covariance forecasts (see Symitsi et al. 2018; Bollerslev, Patton, and Quaedvlieg 2018). We construct the global minimum variance portfolio (GMVP) to bypass the problem of forecasting future returns, as the scope of this article is on covariance forecasting. We calculate the weights of GMVP as

$$\min_{\boldsymbol{w}_t} \boldsymbol{w}_t' \hat{\Sigma}_t \boldsymbol{w}_t \qquad \text{s.t.} \qquad \boldsymbol{w}_t' \mathbf{1} = 1 \tag{11}$$

where $\mathbf{1} = (1, \ldots, 1)' \in \mathbb{R}^N$. The optimal weights of the GMVP are given by $\hat{\boldsymbol{w}}_t = \frac{\hat{\Sigma}_t^{-1} \mathbf{1}}{\mathbf{1}' \hat{\Sigma}_t^{-1} \mathbf{1}}$.

In addition, we consider another portfolio imposed with the **long-only** constraint, denoted as GMVP$^+$, which is a more realistic setting in certain markets. In this case, we are not allowed to short-sell stocks, that is, $w_{i,t} \geq 0, \ \forall i, t$. Since there is no closed-form solution for GMVP$^+$, we use the quadratic programming solver in the built-in Python package CVXPY[20] to compute the optimized portfolio weights. Finally, we compute the portfolio returns as $r_t^{(p)} = \hat{\boldsymbol{w}}_t' \boldsymbol{r}_t$.

Based on the covariance forecasts from the competing models, we obtain multiple corresponding portfolios. We then use the following metrics to evaluate the out-of-sample performance of each portfolio.

[20] https://www.cvxpy.org

- (1) Annualized portfolio standard deviation:

$$\sigma^{(p)} = \sqrt{252 \times \frac{1}{T}\sum_{t=1}^{T}\left(r_t^{(p)} - \bar{r}_t^{(p)}\right)^2}, \tag{12}$$

where T is the length of the out-of-sample period, and $\bar{r}_t^{(p)}$ is the average return of a particular portfolio over the out-of-sample period.

- (2) Average portfolio turnover:

$$TO_t = \sum_{i=1}^{N}\left| w_{i,t+1} - w_{i,t}\frac{1+r_{i,t}}{1+\boldsymbol{w}_t'\boldsymbol{r}_t}\right|, \tag{13}$$

$$\tau^{(p)} = \frac{1}{T-1}\sum_{t=1}^{T-1} TO_t,$$

where TO_t represents the portfolio turnover from day t to day $t+1$. Recall that $w_{i,t}$ (resp. $r_{i,t}$) is the i-th element in $\boldsymbol{w}_t$ (resp. $\boldsymbol{r}_t$).

To further gauge the economic value of covariance forecasts, we also compare the aforementioned portfolios with the naive equally weighted portfolio (denoted as 1/N), a common benchmark in the literature.[21] Table 5 illustrates the annualized standard deviation and average turnover for GMVP and GMVP$^+$ over the out-of-sample period, assuming no transaction cost. The results show that using high-frequency data for forecasting RC leads to superior portfolios compared to the naive 1/N portfolio in terms of standard deviation. Portfolios with graph information, especially GHAR(GL, $\widetilde{L}$) and GHAR(GL, $\widetilde{K}$), are able to generate lower out-of-sample variance compared to the one without any graphs, that is, HAR-DRD, further proving evidence for the existence of network/spillover effects.

With regard to portfolio turnover, except for the 1/N portfolio, the portfolios constructed from models that employ complete graphs for both volatilities and correlations yield the lowest average turnover. This observation is expected to some degree, given that these models incorporate aggregated volatility (or correlation) with equal weights on all individual stocks (correlation pairs). Consequently, the covariances tend to remain relatively consistent over time, resulting in reduced portfolio turnover. Moreover, portfolios based on GL graphs generally exhibit lower turnover relative to the portfolios based on sector graphs, or without graphs.

3.5 Longer Future Horizons

One-day forecasting is not the only time horizon of interest to practitioners. In this section, we consider various prediction horizons, which allows us to examine how the impact of graph information varies on future covariances. To obtain longer horizon forecasts, we adopt the direct approach (in line with Andersen et al. 2003; Bollerslev, Patton, and Quaedvlieg 2018), rather than the iterated approach.[22] Specifically, for the GHAR

[21] DeMiguel, Garlappi, and Uppal (2009) demonstrated that of the fourteen models they evaluate across seven empirical datasets, none is consistently better than the 1/N portfolio in terms of Sharpe ratio, certainty-equivalent return, or turnover.

[22] Bollerslev, Patton, and Quaedvlieg (2018) demonstrated that even minor model mis-specifications will be amplified in the iterated one-day-ahead forecasts, and result in worse estimations than the direct forecast procedure in practice.

Table 5. Out-of-sample portfolio performance over 1-day forecast horizon

	GMVP		GMVP$^+$	
	$\sigma^{(p)}$	$\tau^{(p)}$	$\sigma^{(p)}$	$\tau^{(p)}$
1/N	11.890	0.007	11.890	0.007
HAR-Cholesky	9.948	0.664	9.917	0.411
HAR-DRD	9.918	0.771	9.936	0.511
GHAR(-, $\widetilde{K}$)	9.967	0.730	9.935	0.482
GHAR(-, $\widetilde{L}$)	9.970	0.761	9.951	0.510
GHAR(S, -)	9.906	0.738	10.005	0.503
GHAR(S, $\widetilde{K}$)	9.933	0.675	9.983	0.468
GHAR(S, $\widetilde{L}$)	9.930	0.717	9.993	0.499
GHAR(K, -)	9.857	0.653	9.947	0.432
GHAR(K, $\widetilde{K}$)	9.896	0.590	9.930	0.389
GHAR(K, $\widetilde{L}$)	9.893	0.634	9.945	0.426
GHAR(GL, -)	9.831	0.695	9.911	0.475
GHAR(GL, $\widetilde{K}$)	9.825	0.630	**9.883**	0.435
GHAR(GL, $\widetilde{L}$)	**9.799**	0.675	9.901	0.470

Note: The table reports the out-of-sample performance of GMVP and GMVP$^+$ constructed using the 1-day-ahead covariance forecasts of various models. $\sigma^{(p)}$ is the annualized portfolio standard deviation, and $\tau^{(p)}$ is the average portfolio turnover. The lowest $\sigma^{(p)}$ in each column is indicated in bold.

formulation of volatility in Equation (7), we implement the following model over longer forecast horizons. We apply the same idea to the modeling of correlations, or the vector of Cholesky decomposition of covariances.

$$\begin{aligned}\textbf{GHAR}(A): \quad \boldsymbol{v}_{t:t+h} &= \boldsymbol{\alpha}^{(D)} + \beta_d^{(D)} \boldsymbol{v}_{t-1} + \beta_w^{(D)} \boldsymbol{v}_{t-5:t-2} + \beta_m^{(D)} \boldsymbol{v}_{t-22:t-6} \\ &+ \gamma_d^{(D)} \mathbf{W} \cdot \boldsymbol{v}_{t-1} + \gamma_w^{(D)} \mathbf{W} \cdot \boldsymbol{v}_{t-5:t-2} + \gamma_m^{(D)} \mathbf{W} \cdot \boldsymbol{v}_{t-22:t-6},\end{aligned} \tag{14}$$

where $\boldsymbol{v}_{t:t+h} = \sum_{k=0}^{h} \boldsymbol{v}_{t+k}$, and $h = 4$ and 21 for the weekly and monthly forecasts, respectively.

Table 6 illustrates the statistical performance of GHAR models over longer prediction horizons.[23] We observe that GHAR(GL, $\widetilde{L}$) continuously outperforms the baseline HAR-DRD, with only one exception: the $\mathcal{L}^Q$ loss over the monthly forecasting horizon. Furthermore, the ratios converge to one as the prediction horizon gets longer, as expected. Longer time horizon forecasting is less sensitive to graph information. A possible explanation for this might be that unexpected information (e.g., shocks) of a specific stock is diffused to the broad equity market in several days, rendering less necessary the need to explicitly incorporate graph effects into the model, at longer horizons.

4 Robustness Analysis

In this section, we address potential concerns regarding the robustness of our results. In particular, we aim to answer two important questions related to (i) the stability across different market regimes, and (ii) the measurement errors of volatilities. Due to space considerations and to improve the readability, discussions about other relevant concerns, such as the lagged daily graph effect, alternative model update frequencies, and transformations for volatilities and correlations, are provided in Appendices C, D, and E.

[23] We report the portfolio performance results in Appendix B.

Table 6. Out-of-sample losses over longer forecast horizons

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	Rank	p_{MCS}	Ratio	Rank	p_{MCS}	Ratio	Rank	p_{MCS}
Panel A: 1-Week									
HAR-Cholesky	1.051	13	<0.001	1.058	13	<0.001	1.028	13	<0.001
HAR-DRD	1.000	12	0.006	1.000	12	0.013	1.000	7	0.002
GHAR(-, $\widetilde{K}$)	0.993	11	0.008	0.992	11	0.016	0.997	4	0.492
GHAR(-, $\widetilde{L}$)	0.992	10	0.013	0.991	9	0.025	0.997	3	0.492
GHAR(S, -)	0.991	9	0.008	0.992	10	0.011	1.005	12	<0.001
GHAR(S, $\widetilde{K}$)	0.985	8	0.008	0.985	8	0.016	1.003	10	0.002
GHAR(S, $\widetilde{L}$)	0.984	7	0.015	0.983	7	0.025	1.003	11	0.054
GHAR(K, -)	0.981	6	0.028	0.982	6	0.027	1.002	9	<0.001
GHAR(K, $\widetilde{K}$)	0.976	4	0.030	0.977	4	0.027	0.998	5	0.062
GHAR(K, $\widetilde{L}$)	0.975	3	0.064	0.976	2	0.449	0.998	6	0.104
GHAR(GL, -)	0.980	5	0.030	0.982	5	0.027	1.000	8	<0.001
GHAR(GL, $\widetilde{K}$)	0.975	2	0.064	0.976	3	0.055	0.997	2	0.492
GHAR(GL, $\widetilde{L}$)	0.974	1	1.000	0.975	1	1.000	0.997	1	1.000
Panel B: 1-Month									
HAR-Cholesky	1.108	13	<0.001	1.102	13	<0.001	1.023	13	<0.001
HAR-DRD	1.000	11	0.105	1.000	11	0.114	1.000	3	0.006
GHAR(-, $\widetilde{K}$)	0.994	4	0.480	0.993	4	0.581	0.999	1	1.000
GHAR(-, $\widetilde{L}$)	0.993	3	0.697	0.992	3	0.837	0.999	2	0.006
GHAR(S, -)	0.999	9	0.053	0.999	9	0.055	1.010	11	0.006
GHAR(S, $\widetilde{K}$)	0.995	6	0.314	0.994	6	0.501	1.010	10	0.006
GHAR(S, $\widetilde{L}$)	0.994	5	0.374	0.993	5	0.542	1.011	12	0.006
GHAR(K, -)	1.000	12	0.036	1.000	12	0.036	1.010	9	0.005
GHAR(K, $\widetilde{K}$)	1.000	10	0.065	1.000	10	0.064	1.008	7	0.006
GHAR(K, $\widetilde{L}$)	0.999	8	0.146	0.999	8	0.160	1.008	8	0.006
GHAR(GL, -)	0.996	7	0.322	0.996	7	0.393	1.003	6	0.006
GHAR(GL, $\widetilde{K}$)	0.992	2	0.697	0.991	2	0.837	1.001	4	0.006
GHAR(GL, $\widetilde{L}$)	0.991	1	1.000	0.991	1	1.000	1.001	5	0.006

Note: The table reports the out-of-sample losses of forecasting 1-week-ahead (Panel A) and 1-month-ahead (Panel B) RCs.

4.1 Stability across Market Regimes

The first essential question is whether the model performance varies across different market regimes. To this end, we perform a stratified out-of-sample analysis over two sub-samples: a relatively calm period when the realized variance of S&P500 ETF index is below the 90% quantile of its entire sample distribution, and a turbulent period when this realized variance is above its 90% quantile (Pascalau and Poirier 2021).

Table 7 indicates the relative losses, ranks and p-values of the MCS test for the bottom 90% (low and moderate volatility regimes) and top 10% (high volatility regimes) of market RVs. In general, it appears that GHAR(GL, $\widetilde{L}$) is superior across market regimes over all loss functions in both situations. Interestingly, the larger percentage improvements in terms of $\mathcal{L}^E$ and $\mathcal{L}^F$ stem from the turbulent period, while the opposite holds true for $\mathcal{L}^Q$.

Additionally, it is worth noting that HAR-Cholesky achieves worse predictive performance relative to HAR-DRD during the turbulent period. This finding provides robust evidence in favor of the DRD decomposition for modeling covariance matrices, especially during periods of high volatility, in line with findings reported in Andersen et al. (2006).[24]

[24] Andersen et al. (2006) proposed to characterize the dependencies in volatilities and correlations by regime-switching models.

Table 7. Stratified out-of-sample losses

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}
Panel A: Bottom 90%									
HAR-Cholesky	0.987	7	<0.001	0.988	7	<0.001	1.067	13	<0.001
HAR-DRD	1.000	13	<0.001	1.000	13	<0.001	1.000	9	<0.001
GHAR(-, $\widetilde{K}$)	0.999	12	<0.001	0.998	12	<0.001	0.983	7	<0.001
GHAR(-, $\widetilde{L}$)	0.997	11	<0.001	0.996	11	<0.001	0.980	3	0.729
GHAR(S, -)	0.992	10	<0.001	0.993	10	<0.001	1.003	12	<0.001
GHAR(S, $\widetilde{K}$)	0.992	9	<0.001	0.993	9	<0.001	0.984	8	<0.001
GHAR(S, $\widetilde{L}$)	0.990	8	<0.001	0.990	8	<0.001	0.981	4	0.001
GHAR(K, -)	0.986	3	0.676	0.987	3	0.503	1.002	11	<0.001
GHAR(K, $\widetilde{K}$)	0.986	6	<0.001	0.988	5	<0.001	0.982	5	<0.001
GHAR(K, $\widetilde{L}$)	0.985	2	0.780	0.986	2	0.664	0.979	2	0.965
GHAR(GL, -)	0.986	4	0.651	0.988	6	0.367	1.002	10	<0.001
GHAR(GL, $\widetilde{K}$)	0.986	5	<0.001	0.988	4	<0.001	0.982	6	<0.001
GHAR(GL, $\widetilde{L}$)	0.984	1	1.000	0.986	1	1.000	0.979	1	1.000
Panel B: Top 10%									
HAR-Cholesky	1.119	13	<0.001	1.115	13	<0.001	1.091	13	<0.001
HAR-DRD	1.000	12	0.034	1.000	12	0.035	1.000	12	0.011
GHAR(-, $\widetilde{K}$)	0.983	10	0.071	0.981	10	0.117	0.999	10	0.015
GHAR(-, $\widetilde{L}$)	0.982	9	0.147	0.981	9	0.136	0.998	9	0.015
GHAR(S, -)	0.986	11	0.028	0.986	11	0.029	1.000	11	0.006
GHAR(S, $\widetilde{K}$)	0.970	6	0.116	0.968	5	0.210	0.996	7	0.015
GHAR(S, $\widetilde{L}$)	0.970	5	0.292	0.968	6	0.343	0.996	6	0.015
GHAR(K, -)	0.976	7	0.055	0.976	7	0.057	0.998	8	0.015
GHAR(K, $\widetilde{K}$)	0.961	2	0.991	0.959	1	1.000	0.993	4	0.024
GHAR(K, $\widetilde{L}$)	0.961	4	0.991	0.960	2	0.985	0.992	3	0.050
GHAR(GL, -)	0.976	8	0.071	0.976	8	0.076	0.996	5	0.015
GHAR(GL, $\widetilde{K}$)	0.961	3	0.991	0.960	3	0.985	0.991	2	0.188
GHAR(GL, $\widetilde{K}$)	0.961	1	1.000	0.960	4	0.980	0.991	1	1.000

Note: Stratified losses during trading days with the bottom 90% (Panel A) and top 10% (Panel B) realized variance of S&P500 ETF index.

During turbulent periods, the MCS tests contain nearly all models incorporating graph information, as seen in both Euclidean and Frobenius loss measures. According to Page 6 in Patton (2011), "the average QLIKE loss will be less affected (generally) by the most extreme observations in the sample". A few extreme observations may have an unduly large impact on the outcomes of $\mathcal{L}_t^E$ ($\mathcal{L}_t^F$). The inherent volatility of $\mathcal{L}t^E$ ($\mathcal{L}t^F$) during turbulent times presents a challenge for the MCS test in effectively distinguishing the best model set at a given significance level. On the other hand, for QLIKE, the GL graph for volatilities appears preferable during turbulent periods. This suggests that GL graphs demonstrate higher adaptability to structural changes compared to static graphs, such as sector and complete graphs.

In times of market tranquility, according to $\mathcal{L}_t^E$ ($\mathcal{L}_t^F$), the best set fully excludes models based on a complete adjacency matrix for correlations, and the QLIKE loss-based procedure exclusively favors models incorporating line graph. These findings suggest that during periods of relative market stability, it is more appropriate to consider the connected neighbors rather than the entire market when constructing the graph representation for correlations. The sufficiency of line graph information for this purpose is a topic that lies somewhat beyond the immediate scope of this article and could be an avenue for exploration in future research.[25]

4.2 Measurement Errors of Volatilities

Under specific conditions, the realized variance demonstrates convergence to the true latent volatility as the sampling frequency tends to zero. However, in real-world scenarios, there exists a lower limit on the return horizon that remains practically useful for computing RV. This constraint arises due to discrete observations of returns and market microstructure characteristics, such as price grid discretization and bid-ask spreads, which can disrupt the semi-martingale property at extremely high return frequencies. In any finite sample, RV is inevitably subject to measurement error.

Bollerslev, Patton, and Quaedvlieg (2016) revealed that the beta coefficients in the standard HAR model may be affected by measurement errors in the realized variances. By exploiting the asymptotic theory for high-frequency realized variance estimation, the authors proposed an easy-to-implement model, termed as HARQ, shown in Equation (16). The realized quarticity (RQ) is estimated according to Equation (15), aiming to correct the measurement errors.

$$RQ_{i,t} = \frac{M}{3}\sum_{l=1}^{M} r_{i,t(l)}^4, \quad \boldsymbol{RQ}_{t-1} = (RQ_{1,t}, \ldots, RQ_{N,t})', \tag{15}$$

$$\boldsymbol{v}_t = \boldsymbol{\alpha}^{(Q)} + \beta_d^{(Q)}\boldsymbol{v}_{t-1} + \phi_d^{(Q)}\sqrt{\boldsymbol{RQ}_{t-1}} \circ \boldsymbol{v}_{t-1} + \beta_w^{(Q)}\boldsymbol{v}_{t-5:t-2} + \beta_m^{(Q)}\boldsymbol{v}_{t-22:t-6} + \boldsymbol{u}_t^{(Q)}, \tag{16}$$

where $\circ$ denotes the Hadamard (element-wise) product. Bollerslev, Patton, and Quaedvlieg. (2016) showed that the HARQ model outperforms HAR, when applied to the S&P500 ETF index and individual stocks.

To examine if our results are sensitive to measurement errors of volatilities, we also implement the following models with graph information, denoted as **GHARQ** (Equation (17)). As demonstrated by Bollerslev, Patton, and Quaedvlieg (2018), the measurement errors for the correlations are fairly small; therefore, we follow their setting and ignore the attenuation for the correlations. Finally, as summarized in Table 1, we also consider various graphs for modeling volatilities and correlations.

$$\begin{aligned}\boldsymbol{v}_t = {} & \boldsymbol{\alpha}^{(Q)} + \beta_d^{(Q)}\boldsymbol{v}_{t-1} + \phi_d^{(Q)}\sqrt{\boldsymbol{RQ}_{t-1}} \circ \boldsymbol{v}_{t-1} + \beta_w^{(Q)}\boldsymbol{v}_{t-5:t-2} + \beta_m^{(Q)}\boldsymbol{v}_{t-22:t-6} \\ & + \gamma_d^{(Q)}\mathbf{W}\cdot\boldsymbol{v}_{t-1} + \gamma_w^{(Q)}\mathbf{W}\cdot\boldsymbol{v}_{t-5:t-2} + \gamma_m^{(Q)}\mathbf{W}\cdot\boldsymbol{v}_{t-22:t-6} + \boldsymbol{u}_t^{(Q)}.\end{aligned} \tag{17}$$

Table 8 presents the results obtained from **GHARQ**. We conclude from this table that GHARQ(GL, $\widetilde{L}$) remains the best performing model. This indicates that volatility and correlation graphs indeed provide additional information for covariance forecasting. We also observe (unreported) that HARQ-DRD provides significantly better forecasts than HAR-DRD, consistent with Bollerslev, Patton, and Quaedvlieg (2018).

5 Conclusion

This study investigates whether the graph information can provide additional predictive power when forecasting realized variances and covariances in the U.S. equity market. We propose a new approach that augments the HAR-DRD model by considering graph effects from connected assets. We evaluate the in-sample and out-of-sample contributions of such

[25] We would like to thank an anonymous reviewer for suggesting this analysis.

Table 8. Out-of-sample losses of HARQ and GHARQ

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}
HARQ-DRD	1.000	12	<0.001	1.000	12	0.003	1.000	8	<0.001
GHARQ(-, $\widetilde{K}$)	0.995	10	<0.001	0.994	10	0.003	0.992	6	<0.001
GHARQ(-, $\widetilde{L}$)	0.993	9	0.006	0.992	8	0.016	0.989	4	<0.001
GHARQ(S, -)	0.996	11	0.003	0.997	11	0.003	1.010	11	<0.001
GHARQ(S, $\widetilde{K}$)	0.992	8	0.003	0.993	9	0.003	0.991	5	<0.001
GHARQ(S, $\widetilde{L}$)	0.991	7	0.008	0.991	7	0.016	0.989	3	<0.001
GHARQ(K, -)	0.989	6	0.008	0.990	6	0.016	1.023	12	<0.001
GHARQ(K, $\widetilde{K}$)	0.988	5	0.008	0.989	5	0.016	1.002	10	<0.001
GHARQ(K, $\widetilde{L}$)	0.986	3	0.008	0.987	3	0.016	0.999	7	<0.001
GHARQ(GL, -)	0.987	4	0.008	0.988	4	0.016	1.001	9	<0.001
GHARQ(GL, $\widetilde{K}$)	0.985	2	0.008	0.986	2	0.016	0.984	2	0.006
GHARQ(GL, $\widetilde{L}$)	0.983	1	1.000	0.985	1	1.000	0.982	1	1.000

Note: The table reports the out-of-sample losses of forecasting 1-day-ahead realized covariances of HARQ and GHARQ with different graphs. The baseline here is the HARQ-DRD model.

graph effects in RC matrices forecasting, and observe that the proposed model GHAR(GL, $\widetilde{L}$) achieves the best predictive accuracy consistently across various evaluation metrics. Furthermore, based on the covariance predictions, we construct a global minimum variance portfolio. The portfolio analysis shows that portfolios considering graph effects are able to attain significantly lower out-of-sample variance compared to the traditional models without any graph information or the naive equally weighted portfolio. An assessment of the forecast performance over time shows that the forecast improvements over longer horizons (1-week, 1-month) remain significant, but start to decay as the prediction horizon increases. Furthermore, the robustness tests demonstrate the forecast improvements are observed consistently over the different out-of-sample sub-periods and are insensitive to measurement errors of volatilities. Overall, our results show that graph structures in volatilities and correlations are informative for forecasting RC matrices.

5.1 Future Research Directions

There are several interesting avenues to explore in future research. One direction pertains to other types of graphs, e.g., supply/demand chains (see Tokman et al. 2007; Herskovic et al. 2020), shared analyst coverage (Ali and Hirshleifer 2020), and news co-mentions (Sidorov et al. 2016) that may contain rich structural information about equities. It would be very interesting to compare (and potentially ensemble) the predictive power of different types of graphs.

Another interesting direction pertains to performing a similar analysis as in the present article, but across different markets. For example, according to Rapach et al. (2013), Buncic and Gisler (2016), and Choi, Jiang, and Zhang (2021), U.S.-based equity market information can be used to improve returns/volatility forecasts in a large cross-section of international equity markets. However, there seems to be less information about the impact of international equity markets on the U.S. market (Wilms, Rombouts, and Croux 2021). It would be an interesting study to explore the two-way interplay between U.S. equity market and international equity markets, in a bipartite (and potentially multipartite) graph setting.

Future research directions also encompass enhancing the theoretical underpinnings of our proposed model. This involves drawing inspiration from the literature that bridges the gap between discrete-time series models and continuous-time diffusion processes (Kim and

Wang 2016; Song et al. 2021; Kim 2024; Kim, Shin, and Wang 2022), thereby presenting an exciting avenue to explore the relations between low-frequency time series dynamics and high-frequency-based continuous diffusion processes, particularly within the framework of our graph-based approach. Additionally, our considerations extend to the prospect of establishing asymptotic theorems and consistency results in forthcoming research (Kim and Fan 2019; Shin et al. 2021).

Appendix

Appendix A: Data Statistics

The data used in this article are sourced from LOBSTER, which is subject to licensing restrictions and must be purchased by users, as redistribution is not permitted. The source code is available at: github.com/chaozhang-ox/Graph-based-HAR.

Table A.1 provides summary statistics for the volatility estimates and the corresponding sector of each stock. Figure A.1 reports the average daily realized correlations across the twenty-seven stocks. We calculate the realized correlation matrix of stock returns for each day and then compute an average across all trading days. All cross-correlations are positive and range between 0.17 and 0.63. The correlation matrix displays a block-diagonal structure that corresponds to GICS sector classification, which is in line with previous studies (e.g. Benzaquen et al. 2017).

Table A.1. Summary statistics of realized volatilities

	Mean	Std	Min	25%	50%	75%	Max	Sector
AAPL	2.42	4.12	0.07	0.70	1.25	2.48	60.66	Information Technology
MSFT	1.90	3.03	0.11	0.67	1.09	1.93	49.44	Information Technology
INTC	2.40	3.85	0.14	0.86	1.40	2.47	66.06	Information Technology
CSCO	2.08	3.56	0.14	0.70	1.14	2.10	67.58	Information Technology
CRM	4.23	5.82	0.22	1.45	2.43	4.72	83.92	Information Technology
IBM	1.46	2.85	0.11	0.47	0.75	1.35	44.18	Information Technology
DIS	2.01	3.74	0.12	0.60	1.01	1.90	59.95	Communication Services
VZ	1.48	2.89	0.10	0.50	0.78	1.35	47.47	Communication Services
HD	2.24	4.44	0.15	0.63	1.03	2.02	78.39	Consumer Discretionary
MCD	1.26	2.91	0.08	0.40	0.61	1.14	57.51	Consumer Discretionary
NKE	2.15	3.70	0.14	0.75	1.16	2.06	67.88	Consumer Discretionary
PG	1.09	2.55	0.09	0.38	0.59	0.99	56.07	Consumer Staples
KO	1.07	2.24	0.07	0.37	0.58	1.01	42.69	Consumer Staples
WMT	1.26	2.21	0.11	0.45	0.68	1.19	42.30	Consumer Staples
JNJ	1.01	2.11	0.06	0.35	0.55	0.92	43.93	Health Care
UNH	2.89	5.22	0.16	0.80	1.37	2.62	73.55	Health Care
MRK	1.78	3.13	0.12	0.59	0.94	1.77	52.48	Health Care
AMGN	2.02	2.98	0.16	0.83	1.29	2.17	48.47	Health Care
JPM	3.74	9.05	0.15	0.75	1.36	2.83	204.66	Financials
AXP	3.41	7.51	0.12	0.65	1.16	2.68	141.26	Financials
GS	3.51	8.77	0.19	0.94	1.51	2.82	247.35	Financials
TRV	2.23	5.22	0.11	0.50	0.83	1.80	88.70	Financials
BA	2.96	7.59	0.13	0.80	1.36	2.61	211.14	Industrials
HON	1.98	4.08	0.10	0.53	0.98	1.88	89.63	Industrials
MMM	1.52	2.76	0.08	0.48	0.83	1.50	38.19	Industrials
CAT	2.92	4.65	0.15	0.95	1.59	2.92	60.49	Industrials
CVX	2.16	4.47	0.13	0.62	1.09	2.05	78.48	Energy

Note: The table reports summary statistics for the daily realized volatility and the corresponding sector of twenty-seven stocks in DJIA. The statistics are averaged across each trading day.

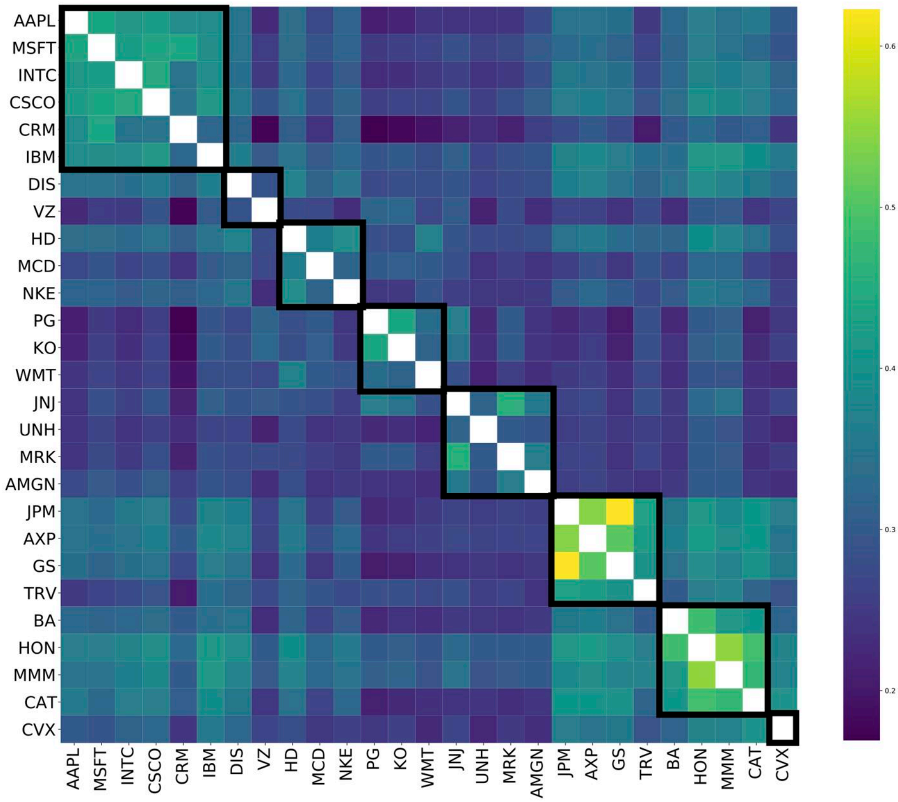


Figure A.1 Realized correlation matrix. The figure shows the daily realized cross-correlations of twenty-seven stocks in DJIA, averaged across each trading day. Stocks are sorted by their industrial sectors, as listed in Appendix Table A.1.

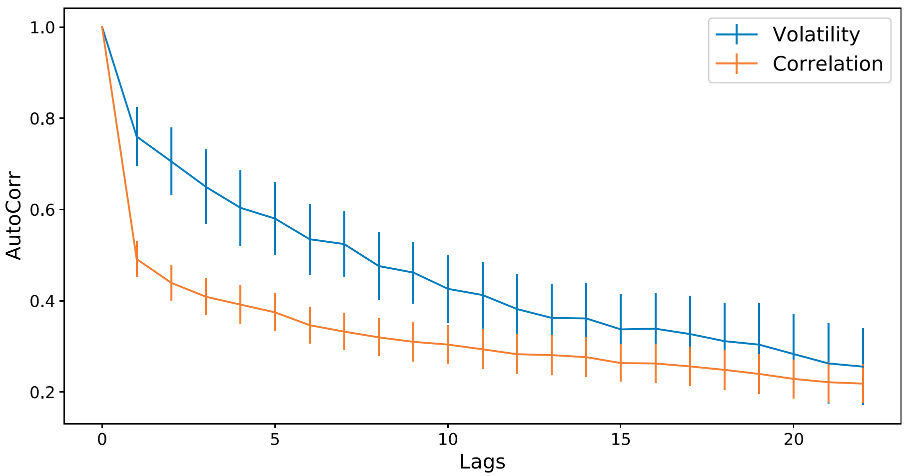


Figure A.2 Autocorrelation of realized volatilities and correlations. The figure provides the standard autocorrelation coefficients and their standard deviations (vertical error bars) of the daily realized volatilities and correlations. The autocorrelation statistics for volatilities (resp. correlations) are averaged across trading days and stocks (resp. stock pairs).

In addition to the usual summary measures, we also report the autocorrelations for realized volatilities and correlations.[26] Figure A.2 reveals that the autocorrelation coefficients of volatilities are consistently larger than those of correlations, directly highlighting the importance of individual models for volatilities and correlations (see Andersen et al. 2006).

Appendix B: Additional Results of Section 3.5

Following Bollerslev, Patton, and Quaedvlieg (2018), the results for GMVP and GMVP$^+$ based on longer horizon forecasting models are averaged across all possible weekly (i.e., Monday-to-Monday, Tuesday-to-Tuesday, etc.) and monthly (i.e., 1st June-to-1st July, 2nd June-to-2nd July, etc.) horizons. Table B.1 details the findings of portfolios rebalanced weekly (Panel A) and monthly (Panel B). To begin, note that modeling volatilities

Table B.1. Out-of-sample portfolio performance over longer forecast horizons

	GMVP		GMVP$^+$	
	$\sigma^{(p)}$	$\tau^{(p)}$	$\sigma^{(p)}$	$\tau^{(p)}$
Panel A: 1-Week				
1/N	10.679	0.016	10.679	0.016
HAR-Cholesky	10.562	0.808	9.641	0.455
HAR-DRD	10.015	0.787	9.536	0.530
GHAR(-, $\widetilde{K}$)	10.000	0.740	9.483	0.488
GHAR(-, $\widetilde{L}$)	10.014	0.768	9.520	0.517
GHAR(S, -)	10.200	0.804	9.703	0.558
GHAR(S, $\widetilde{K}$)	10.177	0.731	9.624	0.510
GHAR(S, $\widetilde{L}$)	10.183	0.770	9.658	0.540
GHAR(K, -)	10.040	0.702	9.700	0.473
GHAR(K, $\widetilde{K}$)	10.033	0.620	9.624	0.414
GHAR(K, $\widetilde{L}$)	10.040	0.664	9.660	0.450
GHAR(GL, -)	9.924	0.754	9.516	0.519
GHAR(GL, $\widetilde{K}$)	**9.884**	0.671	**9.434**	0.463
GHAR(GL, $\widetilde{L}$)	9.892	0.717	9.479	0.498
Panel B: 1-Month				
1/N	10.286	0.035	10.286	0.035
HAR-Cholesky	10.721	1.107	10.130	0.630
HAR-DRD	9.990	0.778	**9.598**	0.558
GHAR(-, $\widetilde{K}$)	9.903	0.814	9.633	0.566
GHAR(-, $\widetilde{L}$)	**9.869**	0.812	9.601	0.578
GHAR(S, -)	10.085	0.840	9.684	0.605
GHAR(S, $\widetilde{K}$)	10.019	0.864	9.714	0.608
GHAR(S, $\widetilde{L}$)	9.989	0.870	9.704	0.622
GHAR(K, -)	10.085	0.795	9.717	0.575
GHAR(K, $\widetilde{K}$)	10.000	0.813	9.727	0.572
GHAR(K, $\widetilde{L}$)	9.969	0.821	9.705	0.588
GHAR(GL, -)	10.258	0.844	9.814	0.615
GHAR(GL, $\widetilde{K}$)	10.194	0.865	9.832	0.615
GHAR(GL, $\widetilde{L}$)	10.178	0.875	9.828	0.633

Note: The table reports the out-of-sample performance of GMVP and GMVP$^+$ constructed using the 1-week-ahead (Panel A) or 1-month-ahead (Panel B) covariance forecasts of various models. $\sigma^{(p)}$ is the annualized portfolio standard deviation, and $\tau^{(p)}$ is the average portfolio turnover. The lowest $\sigma^{(p)}$ attained in each column is indicated in bold.

[26] Hansen and Lunde (2014) proposed a new method to estimate the autocorrelation function of the underlying volatility process using instrumental variable (IV) to mitigate the impact of measurement errors. The authors observed the underlying volatility to be near unit root.

and correlations separately (HAR-DRD) manifests larger economic benefits than the one based on Cholesky decomposition. Compared to the HAR-DRD portfolio, the GHAR portfolios (especially GHAR(GL, $\widetilde{K}$)) achieve better ex-post standard deviations and reduce the turnover over the 1-week horizon. In contrast, the performances of most GHAR models are worse over 1-month horizon. This again provides evidence that the graph effect in the equity market mainly manifests itself at short-term horizons and decays rapidly in time.

Appendix C: Short-Term Graph Effect

The main aim of this section is to assess the performance of models with only the short-term graph effect. In other words, we consider models similar to **GHAR**($\boldsymbol{A}$) (Equation (7)) and **GHAR**($\widetilde{\boldsymbol{A}}$) (Equation (8)), but without the weekly and monthly graph effects. We refer to these models as **GHAR-S**. The results in Table C.1 show that forecast improvements relative to the benchmark HAR-DRD model remain significant even when we only incorporate the short-term graph effect.

Appendix D: Alternative Model Update Frequencies

A potential concern one could raise is that the superior forecasting ability of our GHAR models could be driven to some extent by the low update frequency (every month in our previous analysis) of models. To investigate this possibility, we re-run our analysis and update all models at higher frequencies, specifically daily and weekly. The results in Table D.1 are generally consistent with those from our main analysis.

Appendix E: Transformations for Volatilities and Correlations

In line with Andersen et al. (2006), we now assess the ability of our proposed models under the transformation for volatilities and correlations. Specifically, we apply the log transformation to volatilities (e.g. Andersen et al. 2003; Herskovic et al. 2016; Bucci 2020a; Zhang et al. 2024) and the Fisher transformation to correlations (see Andersen et al. 2006; Hansen, Lunde, and Voev 2014). Table E.1 presents the results for modeling transformed

Table C.1. Out-of-sample losses with the short-term graph effect

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	Rank	p_{MCS}	Ratio	Rank	p_{MCS}	Ratio	Rank	p_{MCS}
HAR-DRD	1.000	12	0.010	1.000	12	0.019	1.000	10	<0.001
GHAR-S(-, $\widetilde{K}$)	0.997	11	0.008	0.997	11	0.013	0.992	8	<0.001
GHAR-S(-, $\widetilde{L}$)	0.996	10	0.010	0.995	9	0.019	0.990	4	0.005
GHAR-S(S, -)	0.994	9	0.010	0.995	10	0.019	1.000	12	<0.001
GHAR-S(S, $\widetilde{K}$)	0.992	8	0.010	0.993	8	0.019	0.992	7	<0.001
GHAR-S(S, $\widetilde{L}$)	0.991	7	0.010	0.991	6	0.019	0.990	3	0.001
GHAR-S(K, -)	0.990	6	0.010	0.991	7	0.019	1.000	9	<0.001
GHAR-S(K, $\widetilde{K}$)	0.989	5	0.010	0.990	5	0.019	0.991	5	<0.001
GHAR-S(K, $\widetilde{L}$)	0.988	3	0.010	0.989	3	0.019	0.989	2	0.387
GHAR-S(GL, -)	0.988	4	0.028	0.990	4	0.026	1.000	11	<0.001
GHAR-S(GL, $\widetilde{K}$)	0.987	2	0.028	0.988	2	0.026	0.991	6	<0.001
GHAR-S(GL, $\widetilde{L}$)	0.986	1	1.000	0.987	1	1.000	0.988	1	1.000

Note: The table reports the out-of-sample losses of the models with the short-term graph effect over the 1-day forecast horizon, averaged over the entire testing sample. $\mathcal{L}^E$ is the Euclidean distance, $\mathcal{L}^F$ is the Frobenius distance, and $\mathcal{L}^Q$ is the Quasi-Likelihood loss function.

Table D.1. Out-of-sample losses under alternative model update frequencies

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}
Panel A: Daily									
HAR-Cholesky	1.026	13	<0.001	1.027	13	<0.001	1.115	13	<0.001
HAR-DRD	1.000	12	0.015	1.000	12	0.023	1.000	9	<0.001
GHAR(-, $\widetilde{K}$)	0.993	10	0.022	0.992	8	0.035	0.986	3	<0.001
GHAR(-, $\widetilde{L}$)	0.991	7	0.035	0.990	7	0.067	0.984	1	1.000
GHAR(S, -)	0.995	11	0.019	0.996	11	0.020	1.007	11	<0.001
GHAR(S, $\widetilde{K}$)	0.989	6	0.030	0.989	6	0.035	0.991	6	<0.001
GHAR(S, $\widetilde{L}$)	0.988	5	0.062	0.988	4	0.110	0.988	5	<0.001
GHAR(K, -)	0.992	9	0.022	0.994	10	0.032	1.015	12	<0.001
GHAR(K, $\widetilde{K}$)	0.987	4	0.035	0.988	5	0.067	0.997	8	<0.001
GHAR(K, $\widetilde{L}$)	0.986	3	0.062	0.987	2	0.140	0.995	7	<0.001
GHAR(GL, -)	0.992	8	0.030	0.993	9	0.033	1.003	10	<0.001
GHAR(GL, $\widetilde{K}$)	0.986	2	0.062	0.987	3	0.110	0.987	4	<0.001
GHAR(GL, $\widetilde{L}$)	0.986	1	1.000	0.986	1	1.000	0.984	2	0.713
Panel B: Weekly									
HAR-Cholesky	1.029	13	<0.001	1.031	13	<0.001	1.104	13	<0.001
HAR-DRD	1.000	12	0.003	1.000	12	0.006	1.000	9	<0.001
GHAR(-, $\widetilde{K}$)	0.991	10	0.008	0.990	8	0.021	0.986	4	<0.001
GHAR(-, $\widetilde{L}$)	0.990	8	0.014	0.988	7	0.037	0.984	2	0.542
GHAR(S, -)	0.994	11	0.006	0.994	11	0.008	1.006	11	<0.001
GHAR(S, $\widetilde{K}$)	0.986	6	0.014	0.986	6	0.033	0.991	8	<0.001
GHAR(S, $\widetilde{L}$)	0.985	5	0.039	0.985	4	0.094	0.988	5	<0.001
GHAR(K, -)	0.990	9	0.010	0.992	10	0.012	1.008	12	<0.001
GHAR(K, $\widetilde{K}$)	0.984	4	0.014	0.985	5	0.037	0.991	7	<0.001
GHAR(K, $\widetilde{L}$)	0.983	3	0.039	0.983	3	0.094	0.988	6	<0.001
GHAR(GL, -)	0.989	7	0.012	0.991	9	0.016	1.002	10	<0.001
GHAR(GL, $\widetilde{K}$)	0.983	2	0.039	0.983	2	0.094	0.986	3	<0.001
GHAR(GL, $\widetilde{L}$)	0.982	1	1.000	0.982	1	1.000	0.983	1	1.000

Note: The table reports the out-of-sample losses of the competitive models when updated every day (Panel A) and every week (Panel B).

Table E.1:. Out-of-sample losses for forecasting transformed volatilities and correlations.

	$\mathcal{L}^E$			$\mathcal{L}^F$			$\mathcal{L}^Q$		
	Ratio	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}	**Ratio**	**Rank**	p_{MCS}
HAR-DRD	1.000	12	0.003	1.000	12	0.002	1.000	12	<0.001
GHAR(-, $\widetilde{K}$)	0.992	11	0.004	0.991	11	0.003	0.979	8	<0.001
GHAR(-, $\widetilde{L}$)	0.991	10	0.004	0.990	10	0.004	0.975	6	<0.001
GHAR(S, -)	0.987	9	0.004	0.986	9	0.002	0.998	11	<0.001
GHAR(S, $\widetilde{K}$)	0.980	8	0.004	0.978	8	0.004	0.976	7	<0.001
GHAR(S, $\widetilde{L}$)	0.979	7	0.004	0.977	7	0.004	0.972	3	<0.001
GHAR(K, -)	0.974	5	0.004	0.973	5	0.004	0.996	9	<0.001
GHAR(K, $\widetilde{K}$)	0.969	2	0.004	0.967	2	0.004	0.973	4	<0.001
GHAR(K, $\widetilde{L}$)	0.968	1	1.000	0.966	1	1.000	0.968	1	1.000
GHAR(GL, -)	0.978	6	0.004	0.977	6	0.004	0.998	10	<0.001
GHAR(GL, $\widetilde{K}$)	0.972	4	0.004	0.970	4	0.004	0.975	5	<0.001
GHAR(GL, $\widetilde{L}$)	0.972	3	0.004	0.970	3	0.004	0.969	2	0.523

Note: The table reports the out-of-sample losses of the models for forecasting transformed volatilities and correlations over the 1-day forecast horizon, averaged over the entire testing sample. $\mathcal{L}^E$ is the Euclidean distance, $\mathcal{L}^F$ is the Frobenius distance, and $\mathcal{L}^Q$ is the Quasi-Likelihood loss function.

volatilities and correlations. As shown, GHAR(K, $\widetilde{L}$) is the best performing model, followed by GHAR(GL, $\widetilde{L}$). The baseline model HAR-DRD is never included in the MCS, demonstrating the importance of graph structural information.

References

Acemoglu, Daron,Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi. 2012. The Network Origins of Aggregate Fluctuations. *Econometrica* 80: 1977–2016.

Aït-Sahalia, Yacine, and Dacheng Xiu. 2016. Increased Correlation among Asset Classes: Are Volatility or Jumps to Blame, or Both? *Journal of Econometrics* 194: 205–219.

Ali, Usman, and David Hirshleifer. 2020. Shared Analyst Coverage: Unifying Momentum Spillover Effects. *Journal of Financial Economics* 136: 649–675.

Andersen, Torben G., Tim Bollerslev, Francis X. Diebold, and Paul Labys. 2003. Modeling and Forecasting Realized Volatility. *Econometrica* 71: 579–625.

Andersen, Torben G., Tim Bollerslev, Peter F. Christoffersen, and Francis X. Diebold. 2006. Volatility and Correlation Forecasting. *Handbook of Economic Forecasting* 1: 777–878.

Andersen, Torben G., Tim Bollerslev, and Nour Meddahi. 2011. Realized Volatility Forecasting and Market Microstructure Noise. *Journal of Econometrics* 160: 220–234.

Baele, Lieven. 2005. Volatility Spillover Effects in European Equity Markets. *Journal of Financial and Quantitative Analysis* 40: 373–401.

Barndorff-Nielsen, Ole E., Peter Reinhard Hansen, Asger Lunde, and Neil Shephard. 2011. Multivariate Realised Kernels: Consistent Positive Semi-Definite Estimators of the Covariation of Equity Prices with Noise and Non-Synchronous Trading. *Journal of Econometrics* 162: 149–169.

Baruník, Jozef, Evžen Kočenda, and Lukáš Vácha. 2016. Asymmetric Connectedness on the US Stock Market: Bad and Good Volatility Spillovers. *Journal of Financial Markets* 27: 55–78.

Bauwens, Luc, and Edoardo Otranto. 2023. Modeling Realized Covariance Matrices: A Class of Hadamard Exponential Models. *Journal of Financial Econometrics* 21: 1376–1401.

Bauwens, Luc, and Yongdeng Xu. 2023. DCC-and DECO-HEAVY: Multivariate GARCH Models Based on Realized Variances and Correlations. *International Journal of Forecasting* 39: 938–955.

Bauwens, Luc, Giuseppe Storti, and Francesco Violante, et al. 2012. Dynamic Conditional Correlation Models for Realized Covariance Matrices. *Core DP* 60: 104–108.

Benzaquen, Michael, Iacopo Mastromatteo, Zoltan Eisler, and Jean-Philippe Bouchaud. 2017. Dissecting Cross-Impact on Stock Markets: An Empirical Analysis. *Journal of Statistical Mechanics: Theory and Experiment* 2017: 023406.

Bollerslev, Tim, Andrew J. Patton, and Rogier Quaedvlieg. 2016. Exploiting the Errors: A Simple Approach for Improved Volatility Forecasting. *Journal of Econometrics* 192: 1–18.

Bollerslev, Tim, Benjamin Hood, John Huss, and Lasse Heje Pedersen. 2018. Risk Everywhere: Modeling and Managing Volatility. *The Review of Financial Studies* 31: 2729–2773.

Bollerslev, Tim, Andrew J. Patton, and Rogier Quaedvlieg. 2018. Modeling and Forecasting (un) Reliable Realized Covariances for More Reliable Financial Decisions. *Journal of Econometrics* 207: 71–91.

Bollerslev, Tim, Marcelo C. Medeiros, Andrew J. Patton, and Rogier Quaedvlieg. 2022. From Zero to Hero: Realized Partial (co) Variances. *Journal of Econometrics* 231: 348–360.

Bonato, Matteo, Massimiliano Caporin, and Angelo Ranaldo. 2013. Risk Spillovers in International Equity Portfolios. *Journal of Empirical Finance* 24: 121–137.

Bucci, Andrea. 2020a. Realized Volatility Forecasting with Neural Networks. *Journal of Financial Econometrics* 18: 502–531.

Bucci, Andrea. 2020b. Cholesky–ANN Models for Predicting Multivariate Realized Volatility. *Journal of Forecasting* 39: 865–876.

Buncic, Daniel, and Katja I. M. Gisler. 2016. Global Equity Market Volatility Spillovers: A Broader Role for the United States. *International Journal of Forecasting* 32: 1317–1339.

Callot, Laurent A. F., Anders B. Kock, and Marcelo C. Medeiros. 2017. Modeling and Forecasting Large Realized Covariance Matrices and Portfolio Choice. *Journal of Applied Econometrics* 32: 140–158.

Chiriac, Roxana, and Valeri Voev. 2011. Modelling and Forecasting Multivariate Realized Volatility. *Journal of Applied Econometrics* 26: 922–947.

Choi, Darwin, Wenxi Jiang, and Chao Zhang. 2021. Alpha Go Everywhere: Machine Learning and International Stock Returns. *Available at SSRN 3489679*

Christensen, Kim, Silja Kinnebrock, and Mark Podolskij. 2010. Pre-Averaging Estimators of the Ex-Post Covariance Matrix in Noisy Diffusion Models with Non-Synchronous Data. *Journal of Econometrics* 159: 116–133.

Corsi, Fulvio. 2009. A Simple Approximate Long-Memory Model of Realized Volatility. *Journal of Financial Econometrics* 7: 174–196.

Costa, Antonio, Paulo Matos, and Cristiano da Silva. 2022. Sectoral Connectedness: New Evidence from US Stock Market during COVID-19 Pandemics. *Finance Research Letters* 45: 102124.

DeMiguel, Victor, Lorenzo Garlappi, and Raman Uppal. 2009. Optimal versus Naive Diversification: How Inefficient is the 1/N Portfolio Strategy? *Review of Financial Studies* 22: 1915–1953.

Diebold, Francis X, and Kamil Yilmaz. 2012. Better to Give than to Receive: Predictive Directional Measurement of Volatility Spillovers. *International Journal of Forecasting* 28: 57–66.

Elliott, Matthew, Benjamin Golub, and Matthew O. Jackson. 2014. Financial Networks and Contagion. *American Economic Review* 104: 3115–3153.

Engle, Robert. 2002. Dynamic Conditional Correlation: A Simple Class of Multivariate Generalized Autoregressive Conditional Heteroskedasticity Models. *Journal of Business & Economic Statistics* 20: 339–350.

Engle, Robert F, and Kenneth F. Kroner. 1995. Multivariate Simultaneous Generalized ARCH. *Econometric Theory* 11: 122–150.

Everett, M. G, and S. P. Borgatti. 1999. The Centrality of Groups and Classes. *The Journal of Mathematical Sociology* 23: 181–201.

Fan, Jianqing, Alex Furger, and Dacheng Xiu. 2016. Incorporating Global Industrial Classification Standard into Portfolio Allocation: A Simple Factor-Based Large Covariance Matrix Estimator with High-Frequency Data. *Journal of Business & Economic Statistics* 34: 489–503.

Fiszeder, Piotr, and Witold Orzeszko. 2021. Covariance Matrix Forecasting Using Support Vector Regression. *Applied Intelligence* 51: 7029–7042.

Forbes, K. J., and Roberto Rigobon. 2002. No Contagion, Only Interdependence: Measuring Stock Market Comovements. *Journal of Finance* 57: 2223–2261.

Friedman, Jerome, Trevor Hastie, and Robert Tibshirani. 2008. Sparse Inverse Covariance Estimation with the Graphical Lasso. *Biostatistics (Oxford, England)* 9: 432–441.

Hallac, David, Youngsuk Park, Stephen Boyd, and Jure Leskovec. 2017. "Network Inference via the Time-Varying Graphical LASSO." *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 205–213. New York: ACM.

Hansen, Peter R., and Asger Lunde. 2014. Estimating the Persistence and the Autocorrelation Function of a Time Series That is Measured with Error. *Econometric Theory* 30: 60–93.

Hansen, Peter R., Asger Lunde, and James M. Nason. 2003. Choosing the Best Volatility Models: The Model Confidence Set Approach. *Oxford Bulletin of Economics and Statistics* 65: 839–861.

Hansen, Peter R., Asger Lunde, and James M. Nason. 2011. The Model Confidence Set. *Econometrica* 79: 453–497.

Hansen, Peter R., Asger Lunde, and Valeri Voev. 2014. Realized Beta GARCH: A Multivariate GARCH Model with Realized Measures of Volatility. *Journal of Applied Econometrics* 29: 774–799.

Hautsch, Nikolaus, Lada M. Kyj, and Peter Malec. 2015. Do High-Frequency Data Improve High-Dimensional Portfolio Allocations? *Journal of Applied Econometrics* 30: 263–290.

Hayashi, Takaki, and Nakahiro Yoshida. 2005. On Covariance Estimation of Non-Synchronously Observed Diffusion Processes. *Bernoulli* 11: 359–379.

Henderson, Harold V, and S. R. Searle. 1979. Vec and Vech Operators for Matrices, with Some Uses in Jacobians and Multivariate Statistics. *Canadian Journal of Statistics* 7: 65–81.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh. 2016. The Common Factor in Idiosyncratic Volatility: Quantitative Asset Pricing Implications. *Journal of Financial Economics* 119: 249–283.

Herskovic, Bernard, Bryan Kelly, Hanno Lustig, and Stijn Van Nieuwerburgh. 2020. Firm Volatility in Granular Networks. *Journal of Political Economy* 128: 4097–4162.

Kim, Donggyu. 2024. Exponential Realized GARCH-Itô Volatility Models. *Econometric Theory* 40: 790–826.

Kim, Donggyu, and Jianqing Fan. 2019. Factor GARCH-Itô Models for High-Frequency Data with Application to Large Volatility Matrix Prediction. *Journal of Econometrics* 208: 395–417.

Kim, Donggyu, and Yazhen Wang. 2016. Unified Discrete-Time and Continuous-Time Models and Statistical Inferences for Merged Low-Frequency and High-Frequency Financial Data. *Journal of Econometrics* 194: 220–230.

Kim, Donggyu, Minseok Shin, and Yazhen Wang. 2022. Overnight GARCH-Itô Volatility Models. *Journal of Business & Economic Statistics, Pages* 41: 1215–1227.

King, Mervyn A, and Sushil Wadhwani. 1990. Transmission of Volatility between Stock Markets. *Review of Financial Studies* 3: 5–33.

Lee, Tae-Hwy, and Xiangdong Long. 2009. Copula-Based Multivariate Garch Model with Uncorrelated Dependent Errors. *Journal of Econometrics* 150: 207–218.

Liu, Lily Y., Andrew J. Patton, and Kevin Sheppard. 2015. Does Anything Beat 5-Minute RV? A Comparison of Realized Measures across Multiple Asset Classes. *Journal of Econometrics* 187: 293–311.

Meinshausen, Nicolai, and Peter Bühlmann. 2006. High-Dimensional Graphs and Variable Selection with the LASSO. *The Annals of Statistics* 34: 1436–1462.

Oh, Dong Hwan, and Andrew J. Patton. 2016. High-Dimensional Copula-Based Distributions with Mixed Frequency Data. *Journal of Econometrics* 193: 349–366.

Pascalau, Razvan, and Ryan Poirier. 2021. Increasing the Information Content of Realized Volatility Forecasts. *Journal of Financial Econometrics*

Patton, Andrew J. 2011. Volatility Forecast Comparison Using Imperfect Volatility Proxies. *Journal of Econometrics* 160: 246–256.

Pourhabibi, Tahereh, Kok-Leong Ong, Booi H. Kam, and Yee Ling Boo. 2020. Fraud Detection: A Systematic Literature Review of Graph-Based Anomaly Detection Approaches. *Decision Support Systems* 133: 113303.

Rapach, David E., Jack K. Strauss, and Guofu Zhou. 2013. International Stock Return Predictability: What is the Role of the United States? *The Journal of Finance* 68: 1633–1662.

Shin, Minseok, Donggyu Kim, Yazhen Wang, and Jianqing Fan. 2021. Factor and idiosyncratic VAR-Itô volatility models for heavy-tailed high-frequency financial data. *arXiv preprint arXiv:2109.05227*

Sidorov, S. P, A. R. Faizliev, V. A. Balash, A. A. Gudkov, A. Z. Chekmareva, and P. K. Anikin. 2016. "Company Co-Mention Network Analysis." *International Conference on Network Analysis*, 341–354. New York: Springer.

Song, Xinyu, Donggyu Kim, Huiling Yuan, Xiangyu Cui, Zhiping Lu, Yong Zhou, and Yazhen Wang. 2021. Volatility Analysis with Realized GARCH-Itô Models. *Journal of Econometrics* 222: 393–410.

Sulem, Déborah, Henry Kenlay, Mihai Cucuringu, and Xiaowen Dong. 2024. Graph Similarity Learning for Change-Point Detection in Dynamic Networks. *Machine Learning* 113: 1–44.

Symitsi, Efthymia, Lazaros Symeonidis, Apostolos Kourtis, and Raphael Markellos. 2018. Covariance Forecasting in Equity Markets. *Journal of Banking & Finance* 96: 153–168.

Tokman, Mert, R. Glenn Richey, Louis D. Marino, and K. Mark Weaver. 2007. Exploration, Exploitation and Satisfaction in Supply Chain Portfolio Strategy. *Journal of Business Logistics* 28: 25–56.

Varneskov, Rasmus, and Valeri Voev. 2013. The Role of Realized Ex-Post Covariance Measures and Dynamic Model Choice on the Quality of Covariance Forecasts. *Journal of Empirical Finance* 20: 83–95.

Vassallo, Danilo, Giuseppe Buccheri, and Fulvio Corsi. 2021. A DCC-Type Approach for Realized Covariance Modeling with Score-Driven Dynamics. *International Journal of Forecasting* 37: 569–586.

Voev, Valeri, and Asger Lunde. 2007. Integrated Covariance Estimation Using High-Frequency Data in the Presence of Noise. *Journal of Financial Econometrics* 5: 68–104.

Wang, Jianian, Sheng Zhang, Yanghua Xiao, and Rui Song. 2022. A Review on Graph Neural Network Methods in Financial Applications. *Journal of Data Science* 20: 111–134.

Wang, Yudong, Zhiyuan Pan, and Chongfeng Wu. 2018. Volatility Spillover from the US to International Stock Markets: A Heterogeneous Volatility Spillover GARCH Model. *Journal of Forecasting* 37: 385–400.

Wilms, Ines, Jeroen Rombouts, and Christophe Croux. 2021. Multivariate Volatility Forecasts for Stock Market Indices. *International Journal of Forecasting* 37: 484–499.

Zhang, Chao, Yihuang Zhang, Mihai Cucuringu, and Zhongmin Qian. 2024. Volatility Forecasting with Machine Learning and Intraday Commonality. *Journal of Financial Econometrics* 22: 492–530.

Zhu, Xuening, Rui Pan, Guodong Li, Yuewen Liu, and Hansheng Wang. 2017. Network Vector Autoregression. *The Annals of Statistics* 45: 1096–1123.

Journal of Financial Econometrics, 2024, 23, 1–32
https://doi.org/10.1093/jjfinec/nbae026
Article